DOCUMENT RESUME

ED 099 672 CE 002 792

AUTHOR	Siegel, Arthur I.; And Others
TITLE	A Survey of Student Measurement and Course Evaluation Procedures Within the Air Training Command.
INSTITUTION	Applied Psychological Services, Inc., Wayne, Pa.
SPONS AGENCY	Air Force Human Resources Lab., Brooks AFB, Texas.; Air Force Human Resources Lab., Lowry AFB, Colo. Technical Training Div.
REPORT NO	AFHRL-TR-74-5
PUB DATE	Jul 74
NOTE	117p.
EDRS PRICE	MF-$0.75 HC-$5.40 PLUS POSTAGE
DESCRIPTORS	*Course Evaluation; Educational Assessment; Educational Development; Educational Diagnosis; Evaluation Methods; *Instructional Systems; Interviews; *Measurement Techniques; Military Personnel; *Surveys; Systems Approach; Systems Development; Tables (Data); Technical Education; Testing; *Training
IDENTIFIERS	Air Force

ABSTRACT

The methods, procedures, and results of an interview survey into reactions of user personnel to the Air Force Instructional System Development (ISD) system are reported. Training evaluation, student measurement, and training manager/instructor personnel were interviewed. All groups regarded the ISD procedures favorably, but some problems associated with the ISD procedures were identified. Recommendations for ISD program improvement are presented. The results of an interview survey of the Air Training Command to clarify issues relating to training evaluation and student measurement are reported. The survey was also concerned with a general assessment of the use and application of the Instructional System Development technique. A total of 139 training evaluation, student measurement and training manager/instructor personnel were interviewed. Generally favorable attitudes were reported regarding the use and application of the ISD system, the student measurement techniques, and the course review techniques. Problem areas were identified and specific recommendations for improvement are presented. (Author)

ED 099572

AIR FORCE HUMAN RESOURCES LABORATORY

AFHRL-TR-74-5

A SURVEY OF STUDENT MEASUREMENT AND COURSE EVALUATION PROCEDURES WITHIN THE AIR TRAINING COMMAND

By

Arthur I. Siegel
Philip J. Federman
Applied Psychological Services, Inc.
Wayne, Pennsylvania 19087

Wayne S. Sellman

TECHNICAL TRAINING DIVISION
Lowry Air Force Base, Colorado 80230

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

July 1974

Approved for public release; distribution unlimited.

OE003092

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235

NOTICE

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This final report was submitted by Applied Psychological Services, Inc., Wayne, Pennsylvania 19087, under contract F41609-71-C-0025, project 1121, with the Technical Training Division, Air Force Human Resources Laboratory, Lowry Air Force Base, Colorado 80230. Captain Gary M. Miller, Instructional Technology Branch, was the contract monitor.

This report has been reviewed and cleared for open publication and/or public release by the appropriate Office of Information (OI) in accordance with AFR 190-17 and DoDD 5230.9. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved.

MARTY R. ROCKWAY, Technical Director
Technical Training Division

Approved for publication.

HAROLD E. FISCHER, Colonel, USAF
Commander

BEST COPY AVAILABLE

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFHRL-TR-74-5	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A SURVEY OF STUDENT MEASUREMENT AND COURSE EVALUATION PROCEDURES WITHIN THE AIR TRAINING COMMAND		5. TYPE OF REPORT & PERIOD COVERED Final 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Arthur I. Siegel Philip J. Federman Wayne S. Sellman		8. CONTRACT OR GRANT NUMBER(s) F41609-71-C-0025
9. PERFORMING ORGANIZATION NAME AND ADDRESS Applied Psychological Services, Inc. Wayne, Pennsylvania 19087		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 11210304
11. CONTROLLING OFFICE NAME AND ADDRESS Hq Air Force Human Resources Laboratory Brooks Air Force Base, Texas 78235		12. REPORT DATE July 1974 13. NUMBER OF PAGES 114
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Technical Training Division Air Force Human Resources Laboratory Lowry Air Force Base, Colorado 80230		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

technical training
education
testing
student measurement
training analysis
learning assessment
individual differences
course evaluation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The methods, procedures, and results of an interview survey into reactions of user personnel to the Air Force Instructional System Development (ISD) system are reported. Training evaluation, student measurement, and training manager/instructor personnel were interviewed. All groups regarded the ISD procedures favorably, but some problems associated with the ISD procedures were identified. Recommendations for ISD program improvement are presented. The results of an interview survey of the Air Training Command to clarify issues relating to training evaluation and student measurement are reported. The survey was also concerned with a general assessment of the use and application of the Instructional System Development technique. A total of 139 training evaluation, student measurement and training manager/instructor personnel were interviewed. Generally favorable attitudes were

BEST COPY AVAILABLE

Item 20 (Continued)

reported regarding the use and application of the ISD system, the student measurement techniques, and the course review techniques. Problem areas were identified and specific recommendations for improvement are presented.

BEST COPY AVAILABLE

SUMMARY

Problem

The Human Resources Laboratory initiated a three pronged program for the purpose of clarifying issues relating to training evaluation and student measurement. The field survey of the Air Training Command, reported herein, represents one aspect of this work. The other two aspects of the total program involved: (1) a literature review of the progress made in the training evaluation and student measurement technology over the period of 1950-1970, and (2) preparation of a user-oriented guide to training evaluation and student measurement methods and procedures (AFHRL-TR-72-15). The purposes of the field survey were to obtain data that would help to set into perspective current Air Force training evaluation and student measurement practices and to provide insight into areas needing improvement. Additionally, the Air Training Command survey was concerned with a general assessment of the use and application of the Instructional System Development (ISD) technique.

Approach

The semi-structured interview was employed to collect information at various ATC centers. Separate interviews were developed for each of the four training specialties included in the sample. The specialties sampled were ISD, training evaluation, student measurement, and training managers/instructors. The sample of respondents was drawn from five Air Force Technical Training Centers (Lackland, Sheppard, Keesler, Chanute, and Lowry). The sample was composed of 139 respondents, divided over the four specialties such that 24 percent of the sample were ISD specialists, 15 percent were training evaluation specialists, 22 percent were student measurement specialists, and 39 percent were training managers/instructors.

Results

Each of the four groups thought the ISD procedures achieved a different success level in terms of realizing their general purpose of graduating students who meet job performance requirements, although all groups regarded the ISD procedures favorably. In spite of the generally favorable attitude toward ISD, a number of problems with the ISD procedures were identified by the respondents. These problems were concerned with the need for: (1) an information dissemination program to clarify misconceptions about the ISD system, (2) streamlined ISD procedures, (3) evaluation of each ISD step, and (4) more thorough application of the ISD procedures.

BEST COPY AVAILABLE

The general consensus toward the student measurement techniques employed was also favorable, and the criterion referenced checklist was especially believed to be valuable. However, there was evidence that: (1) some of the written tests employed may lack reliability, (2) standards for test update lack adequacy, (3) parallel test forms may be lacking where they are required, and (4) increased consideration of the confidence and hierarchical testing approaches is warranted.

The principal course review techniques employed were graduate, supervisor, and student critiques. These techniques were regarded as economical and productive of valuable insights into areas for needed course revision. Criticisms of these techniques were related to their lack of objectivity and the interpretation of the data they yield by persons who are unfamiliar with the total subject matter content or with the context in which the recommendations are set.

Conclusions

Among the recommendations for program improvement are: (1) development of an information dissemination program designed to clarify misconceptions about the ISD system and to orient training personnel in the attributes and goals of the ISD program, (2) increased training for personnel involved in ISD use and application (which may also result in a decrease in the time required to meet ISD requirements), (3) development of job aids and work methods, along with an increase in communications efficiency for the purpose of "streamlining" the ISD procedures, (4) evaluation of the proficiency with which each ISD step is performed, (5) establishment of guidelines for updating tests, (6) use of alternate or parallel tests where required, (7) determination of the predictive validity of the tests used, (8) allowing sufficient testing time for the purposes on hand, and (9) development of course review techniques that can yield longitudinal, objective, and quantitative information.

BEST COPY AVAILABLE

TABLE OF CONTENTS

	Page
CHAPTER I - INTRODUCTION	7
Advantages of Systems Approach	9
Purpose of Present Study	10
Collection and Analysis of Field Evaluation Data	11
CHAPTER II - METHOD	16
Development of Interview	16
Initial Assembly of Interview and Pretest	17
Final Interviews	18
Reliability of Interview	22
Interviewer Training	23
Sample	23
CHAPTER III - RESULTS	26
Description of Interviewees	26
Instructional System Development Specialists	28
Emphasis on Each ISD Step	29
Success of Each ISD Step	30
Why Steps 1 and 5 Were Helpful	31
General Opinions	32
Problems Inherent in ISD Technique	33
Training Evaluation Specialists	34
ISD Evaluation	34
Problems Inherent in ISD Technique	35
Training Evaluation	35
Course Review Techniques	36
Identification of Areas of Over and Undertraining	37
Specific Use of Course Reviews	38
Field Evaluation	38

BEST COPY AVAILABLE

TABLE OF CONTENTS (cont.)

	Page
Student Measurement Specialists	40
ISD Evaluation	40
Student Measurement	42
End of Block Written Tests	43
Performance Tests	43
Criterion Checklists	44
Needed Student Measurement Techniques	46
Feedback to Course Developers	46
Review and Revision of Student Measurement Techniques	47
Student Critiques	48
Training Managers/Instructors	48
ISD Evaluation	48
Best Aspects of ISD Method	50
Worst Aspects of ISD Method	52
Evaluation of ISD Program	54
Course Review	54
Field Evaluation	56
Student Measurement	59
Student Critique	60
Evaluation of Instructional Effectiveness	60
General-Nonspecific Responses	60
Importance of ISD Steps	62
Emphasis on ISD Steps	64
Reliability of Interviewee Responses	65
ISD Evaluation as a Function of Location and Experience	67
CHAPTER IV - DISCUSSION	69
Evaluation of ISD	69
Problems Within the ISD Methods	71
Student Measurement Techniques	73
Course Evaluation Techniques	74
Postscript	75
REFERENCES	76
APPENDIX A	77

BEST COPY AVAILABLE

LIST OF TABLES

Table		Page
1	Sample of Respondents.	25
2	Amount of Emphasis Placed on ISD Steps in Course Development	29
3	Amount of Success of ISD Steps 1 and 5	31
4	Results of Evaluation of the ISD Program by ISD Specialists.	32
5	Results of Evaluation of the ISD Program by Training Evaluation Specialists	34
6	Evaluation of the Course Review Techniques Used.	36
7	Evaluation of the Field Evaluation Techniques Used	39
8	Results of Evaluation of the ISD Program by Student Measurement Specialists.	41
9	Student Measurement Techniques Employed.	42
10	Results of Evaluation of ISD Program by Training Managers/ Instructors.	49
11	Positive Adjestives Selected to Describe the ISD Model . . .	50
12	Rank Order of Best Aspects of the ISD Model.	51
13	Negative Adjectives Selected to Describe the ISD Model . . .	52
14	Rank Order of Worst Aspects of the ISD Model	53
15	Evaluation of the Course Review Techniques Used.	55
16	Evaluation of the Field Evaluation Techniques Used	57
17	Mean Rankings of ISD Steps on Importance in Course Development.	63
18	Scale Values of the Emphasis the ISD Steps Received in Course Development	64
19	Reliability Estimates of Interviewee Responses	66
20	Summary of Analysis of Variance for Location and Experience.	68

BEST COPY AVAILABLE

LIST OF FIGURES

Figure		Page
1	Flow diagram for the Instructional System Development model.	8
2	Definitions of the five Instructional System Development steps.	19
3	Mean time in current assignment for sampled interviewees in four training specialties.	27
4	Percentage of each subsample possessing familiarity with courses developed in accordance with ISD principles. .	28
5	Mean magnitude estimates by training specialty (0-100) of ISD for achieving its goals	70

CHAPTER I

INTRODUCTION

The systems approach to training, although not a recent innovation, has received increased emphasis over the recent years. AF Manual 50-2, Instructional System Development, in discussing the background and history of the systems approach to training, points out that the systems approach involves both curriculum planning (in a logical and organized manner, i.e., systematically) and identifying objectives that are to be achieved.

The formal process used by the Air Force in developing training programs within the context of the systems approach, is called Instructional System Development, or ISD. The ISD procedure is composed of five interacting steps, as shown in Figure 1. As also indicated in Figure 1, each step supplies the information needed to accomplish the next step, although portions of various steps may often be accomplished simultaneously.

The five ISD steps are:

1. Analyze System Requirements--identify tasks to be performed within the overall operational environment.

2. Define Education or Training Requirements--determine the tasks that require instruction, the level of proficiency to be developed in students, and the resources needed to conduct the instruction.

3. Develop Objectives and Tests--identify behaviors (attitudes, knowledges, and skills) required for successful job performance and construct criterion and enabling objectives, as well as achievement tests.

4. Plan, Develop, and Validate Instruction--select instructional methods, media, and equipment that best satisfy learning objectives; determine the sequencing of the instructional material; validate the instructional materials to prove that they teach what they are designed to teach and to insure that all elements of the instructional system function effectively in achieving stated objectives.

BEST COPY AVAILABLE

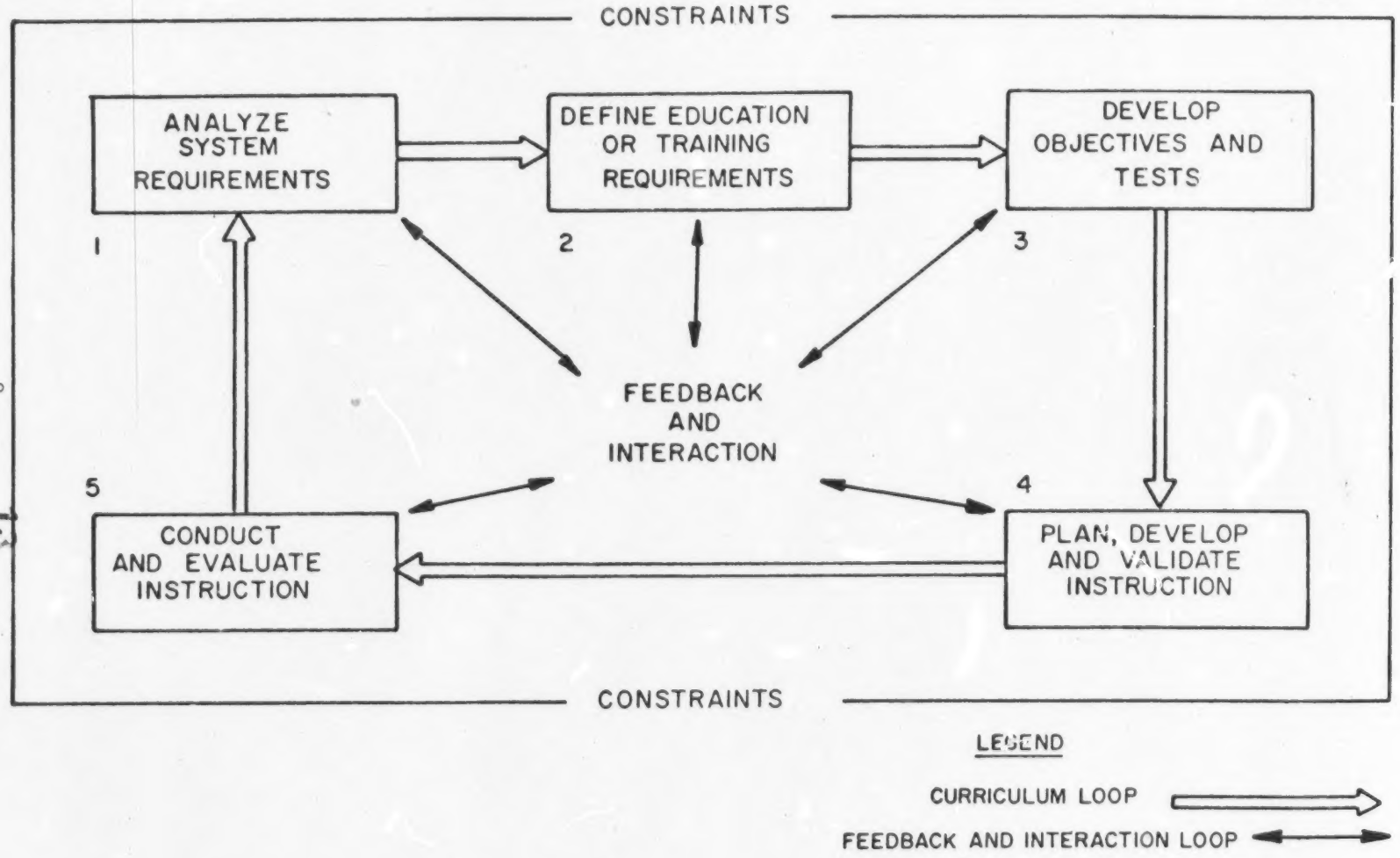


Figure 1. Flow diagram for the instructional system development model

BEST COPY AVAILABLE

5. Conduct and Evaluate Instruction--identify problem areas and corrective actions needed in order to satisfy the requirements of the operating commands.

Advantages of Systems Approach

The systems approach has several advantages, insofar as training program development is concerned. The advantages to be gained by using a systems approach in training program development are:

Comprehensiveness--It ensures that all the elements necessary for on-the-job performance are included in the training.

Job Relatedness--It includes only those elements that pertain directly to successful job performance as determined by an analysis of the system requirements.

Appropriateness--It emphasizes job behaviors which must be demonstrated on the job and the prerequisite skills and knowledges that satisfy job performance requirements. It also specifies the conditions under which the behaviors will take place and the minimum standard of acceptable performance.

Practicality--It includes the concept of cost-effectiveness so that comparisons of desirable features and alternatives are completed.

Flexibility--It includes posttraining course review and student performance analysis. This leads to periodic redesign, modification, and update of the training program.

Validity--The total process ensures that the instruction is related to the training objectives and the performance requirements.

BEST COPY AVAILABLE

Purpose of Present Study

In the fifth step of the ISD model, the actual production of trained personnel is accomplished, and the instructional program is continuously refined and updated in order to ensure that the training program continues to accomplish its goal of producing qualified personnel who meet the requirements of the operating commands. Clearly, effective techniques for the evaluation of training programs and for the measurement of student performance, both during and after training are important aspects of this step, and to the ISD procedure as a whole. In this sense, training evaluation constitutes the quality control aspect of the training system development cycle. Training evaluation, although it might occur in one specific phase of training, provides feedback to all phases of the training program from the very initial analyses of system requirements to the actual conduct of the training. The feedback data are required for training improvement, determining if the training meets its objectives, and comparing cost-effectiveness of different methods and approaches.

Similarly, student measurement yields information which allows the trainer to know whether or not students are meeting course objectives, and if they need special assistance. Student measurement results also find application in the solution to problems related to improved student selection and classification, graduate assignment, and advanced training.

Although some type of training evaluation and student measurement (however formal or informal) takes place in most training programs, there is no universally accepted single best training evaluation or student measurement method. There are many issues involved in training evaluation and student measurement, and most workers in the field seem to agree that training evaluation and student measurement are multidimensional in nature.

In order to clarify issues related to training evaluation and student measurement in the Air Force, the Air Force Human Resources Laboratory initiated a three-pronged program through a contract with the Applied Psychological Services.

Initially, a literature review was conducted. The literature review was concerned with both training evaluation and student measurement. To ensure complete coverage of the relevant fields, searches were completed into the literature covering the period 1950 to 1970, including technical reports, journal articles, books, and annual reviews. The results of this literature review are documented in the report,

BEST COPY AVAILABLE

Training Evaluation and Student Achievement Measurement: A Review of the Literature (AFHRL-TR-72-3). The report documents and describes progress in and the current state of the technology in the content area involved. Current issues and problems are clarified and comments made on the offered solutions.

A second aspect of the approach to ensuring improved training evaluation and student measurement techniques in the Air Force involved the development of a report which deals specifically with training evaluation and student measurement methods and procedures. It summarizes and integrates the evaluation technology state-of-the-art for direct application to Air Force systems. Advantages and disadvantages of various techniques are discussed and, where appropriate, recommendations are made concerning the application of the procedures to specific training evaluation situations. This report is published under the title, Some Techniques for the Evaluation of Technical Training Courses and Students (AFHRL-TR-72-15).

Finally, to achieve the goals of the total study, a field survey of the Air Force Technical Training Centers was performed. The purposes of the field work were to survey and review the training evaluation and student measurement practices at the various training centers. The methods, procedures, and results of this field survey are the concern of the present report. It was anticipated that the findings of such a survey would help to set into perspective current Air Force training evaluation and student measurement practices and provide insight into areas for needed improvement. As a corollary, a general assessment of the use and application of the ISD technique was also included within the field survey.

Collection and Analysis of Field Evaluation Data

Field evaluations are designed to obtain information that can be used by training administrators for improving and updating training. By knowing what operational needs graduates meet and where graduates fail to satisfy field requirements, one can determine needed areas for course revision. Such evaluations are usually conducted after graduates have been on the job for which they were trained for a period of time. There are several field evaluative techniques: graduate and supervisory questionnaires, graduate and supervisory interviews, and evaluations of the graduates' knowledges, skills, and attitudes through tests and rating methods. Each possesses its own advantages and disadvantages.

BEST COPY AVAILABLE

The questionnaire, as a self-administrating instrument, can be mailed to course graduates and/or their supervisors and similarly returned to a central processing office.

Several areas can be investigated through questionnaire methods. One of these is the affective or emotional reactions of graduates to the training program. Another is their specific thoughts and suggestions regarding training program improvement, i.e., training areas of needed additional emphasis or deemphasis on the basis of the graduate's experience on the job.

There are many different types of questionnaire item that can be constructed for obtaining evaluative information. The Likert type item asks the extent of agreement or disagreement with a particular statement. For example, "Do you (strongly agree), (agree), (mildly agree), (mildly disagree), or (strongly disagree) that the training you received prepared you adequately for your present job." Another item type is the semantic differential item. Here, bipolar adjectives are used to define the endpoints of semantic dimensions. An example of a semantic differential item is: "Place a checkmark anywhere along the line, between the pair of words, to indicate your feelings about the training program."

Another popular questionnaire item type is sentence completion. The stem of the sentence, which is provided to the respondent, channels his thoughts in a particular direction. For instance, "One of the things that is done differently on the job from the way it was taught in school is _____."

Multiple choice questions frequently appear in questionnaires. If the response category "other" is included along with the response categories, the changes of losing data are decreased. An example of a multiple choice item is:

> Check the one area of training which was most irrelevant to your present job:
>
> _____1. theory of electronics
> _____2. use of test equipment
> _____3. component identification
> _____4. circuit design and troubleshooting
> _____5. other, specify_____

BEST COPY AVAILABLE

Forced-choice items consist of a group of statements which are matched on favorability. In a two-choice question, one of the items will differentiate in terms of one characteristic, while the other item will differentiate in terms of another characteristic. The statement chose by the respondent reflects his preference. Forced-choice items can have three, four, or five choices as well. One variation of this item type is to have two desirable statements and two undesirable statements, with the respondent selecting the one which most represents his point of view and the one which least represents his point of view. This technique is held to eliminate much response dishonesty. A forced-choice item might appear as follows:

> Place an 'M' next to the option that best describes your feelings to the set. Place an "L" next to the one which least describes your feelings:
>
> a. The course prepared me for adequate function on the job.
>
> b. The tools and equipment in the school aren't much like those on the job.
>
> c. The laboratory exercises were much like the work we do on the job.
>
> d. Much of what we do on the job is different from the way we were taught to do it in school.

The second major technique used for collecting field data for training evaluative purposes, the interview, consists of three basic types: unsystematic, systematic, and standardized. In the unsystematic interview, usually no particular sequence is followed, although the interviewer may have a set of questions he developed in advance. This approach can prove to be highly unreliable and invalid. In the systematic interview, the interviewer has a plan and format. However, within the plan the interviewer is allowed the flexibility to probe where he thinks it is necessary. The standardized interview is characterized by a series of questions that are asked by all interviewers of all interviewees in the same sequence and with the same wording. The technique eliminates the interviewer as a source of response variability.

Finally, training evaluation data can be collected in the field through various testing and rating procedures. Performance testing, in which the performance of the graduate on various job aspects is observed and scored, represents the procedure used most widely.

BEST COPY AVAILABLE

The scoring is based on adherence to correct procedures, adherence of a final product to prescribed standards, care and use of tools, observance to safety precautions, and time for task completion. Typically, the scoring form contains all the sequential elements, which must be performed correctly, in a task. The scorer observes the examinee's performance on the task and scores each subelement accordingly. A final score, the sum of all subscores, is usually obtained from this type of test. Analysis of the test scores yields the data required for training program improvement.

The proving ground represents an advanced adaptation of the performance test situation. Here, the graduate is cycled through all the tasks of a job in a short period of time. As he performs on each task, the graduate is evaluated. He, in turn, evaluates the training he received in relation to the task.

It is possible to build criterion referenced methods into the supervisory questionnaire or the performance testing field evaluation paradigm. When a graduate is evaluated through a criterion referenced technique, the graduate's performance is compared with an absolute standard to determine if he attained acceptable (criterion) performance. The absolute standard may be a performance test item of a given level of difficulty or achievement of an absolute score. The characteristics of criterion referenced scores are that they:

1. indicate the degree of competence attained, independent of the performance of others
2. measure performance with regard to specified absolute standards (criteria) of attainment
3. minimize individual differences
4. consider variability irrelevant

Generally, the criterion referenced score indicates how well a person performs with regard to a specified standard of behavior. Criterion objectives, which are expressed in terms of specific behaviors, must be identified. The criterion objectives specify the operations and knowledges the graduate must demonstrate to satisfy a job performance requirements. Individual differences among examinees are considered irrelevant, since the examinee is graded against a single external standard rather than against other job incumbents. Criterion referenced measurement programs usually use S (satisfactory) and U (unsatisfactory) grades.

BEST COPY AVAILABLE

The criterion referenced approach is desirable in those situations where an indication is wanted of absolute proficiency or ability of the graduate to perform on a particular job. Since an important concept underlying technical training is that of training students to a defined standard as represented by the job requirements, it appears that the criterion referenced approach can be most appropriate relative to skill training.

Several sequential steps are involved in the development and administration of criterion referenced instruments.

1. Following a job analysis, the criterion job behaviors (tasks) are identified on which a graduate must be able to demonstrate competency.

2. Tests are developed, one for each criterion behavior (task), to measure competency.

3. "Minimally acceptable" and "desirable" cut scores are established for each test.

4. The tests are administered and the graduates are evaluated relative to their performance as compared with the cut scores.

If a supervisory evaluation rather than a performance testing approach is followed, then rating scales are developed for each criterion behavior. The supervisor of each person being evaluated rates the performance of each graduate in relationship to each of the cut points.

BEST COPY AVAILABLE

CHAPTER II

METHOD

The semi-structured interview method was employed to collect the information required from the field survey. The semi-structured interview has several advantages over other methods of collecting field data. This technique ensures that all relevant areas are covered by appropriate questions, but permits retaining the flavor and details of responses by recording them in the interviewee's own words. The method also allows for probing by the interviewer in order to gain more insight into areas mentioned by the interviewee. Whenever the interviewer considers it to be necessary, the interview procedure can be varied to meet individual cases.

The face-to-face nature of this type of interview ensures that the interviewee will devote his attention to the questions asked, and that he will respond. The interviewee is likely to feel quite comfortable and free from stress with this technique and, accordingly, is likely to be quite willing to amplify his responses when necessary.

As compared with an interview, a printed questionnaire is less likely to elicit frank statements. The semi-structured interview, with its informal atmosphere, is most likely to lead to a situation in which the interviewee will be candid, and the lattitude allowed the interviewer can be used for establishing cooperation and rapport.

Development of Interview

Prior to developing the interview, an attempt was made to acquire a better insight into the requirements of the training evaluation, student measurement, and the ISD program as they apply to the Air Technical Training Command. Various Air Force manuals were reviewed. This review included, but was not limited to, the following: AFM 50-9--Principles and Techniques of Instruction, AFM 50-2--Instructional System Development (ISD), ATCR 52-1--Training Evaluation and Course Reviews, ATCR 52-3--Measurement, ATCR 52-16--Student Record of Training, ATCR 52-26--Student Scheduling and Administration, ATCR 52-29--Student Critique Program, ATCR 52-18--Training Materials, and ATCR 50-4--Instructor Supervision. Other documents, mainly released as project reports by various ATC centers, were also reviewed. Further insights into the requirements were gained through conferences and discussions with personnel from the project sponsoring staff and the ATC staff.

BEST COPY AVAILABLE

In developing the interview, the areas of interest, i.e., those areas into which the interview would probe, were first identified. These areas were:

- overall evaluation of the ISD program
- evaluation of the specific ISD steps
- general opinions about ISD and its application
- effectiveness of the ISD steps for achieving their purposes
- attitudes toward ISD
- applications of the ISD principles
- specific problems inherent in ISD
- methods for improving ISD
- how ISD was used and how it was helpful
- evaluation of the training evaluation techniques used
- problems associated with the training evaluation techniques employed
- how information from the training evaluations was fed back to the user
- how the user implemented this information
- evaluation of the techniques used to measure student achievement
- problems associated with the student measurement techniques used
- how results of the student measurements are used

Initial Assembly of Interview and Pretest

The interview, in its original form, consisted of over 100 items and included several different data collection techniques. These included the critical incident approach, rating scales, rank ordering, magnitude estimations, and open-ended questions.

This interview was pretested on a sample of seven specialists working at the 3415th TS, LTTC, Lowry AFB, Colorado, in specialties which were to be included in the eventual sample. The pretesting indicated one major required change, aside from the typical editing and rephrasing of questions for the purpose of enhancing clarity and understanding.

BEST COPY AVAILABLE

Since the interview inquired into three major content areas (ISD, training evaluation, and student measurement) questions were asked of the pretest respondents in each of these areas. Yet, those respondents who were questioned on issues outside their immediate concern (either ISD, or training evaluation, or student measurement) were often unable to provide knowledgeable responses. Or, if responses were offered, their accuracy or meaningfulness was suspect. In addition, the length of the interview was unwieldy--it required in excess of two hours.

As a result of the pretest, the original interview was reorganized into five separate interviews. Thus, respondents would receive questions that were specific to their background and experience. The fractionation of the original interview into five separate interviews reduced the time required to complete an interview considerably, although there were many questions that appeared on more than one interview form. Whereas originally over two hours were required for the interview, now the interviews could be completed in one-half hour to one and one-quarter hours, depending on the interview form. The different training specialties into which the interview was divided were training evaluation, student measurement, training managers/instructors, ISD, and headquarters.

Final Interviews

The interview instruments utilized are presented in the Appendix to this report and are described below.

At the start of the interview, but after the initial introduction, explanation of the study, and the purposes of the interview, the interviewee was provided with two handouts. These were the flow diagram of the ISD model, as presented in Figure 1, and a set of definitions of the five ISD steps. The ISD definitions are presented as Figure 2. These were provided so that if the interviewee wished to refamiliarize himself with the concepts, he could, and those interviewees whose knowledge and familiarization with ISD was spotty could refer to the handouts whenever necessary during the interview.

Certain questions in the interview required the interviewee to refer to specific definitions or scales of values. Wherever this was required, handout cards were displayed. These are contained at the end of each interview set in Appendix A.

BEST COPY AVAILABLE

ISD DEFINITIONS

Analyze System Requirements (Step 1): Familiarization with the operational system, the major components and the subsystems, and understanding the system's mission. Identification of the job(s) to be performed; placement of each job in perspective with objectives, requirements, and the overall environment of the operational system.

Define Education or Training Requirements (Step 2): Generation of an inventory of job performance requirements and the required levels of performance. Determination of the following: number of personnel to be qualified and methods for acquiring them, time required to develop the instruction, criticality and learning difficulty of tasks, availability and qualifications of personnel and resources needed to develop and implement the instructional system, and leadtime for development and installation.

Develop Objectives and Tests (Step 3): Identification of knowledges, skills, and attitudes required for successful performance. Development and construction of criterion and enabling objectives and tests to evaluate student achievement and progress and to indicate effectiveness of the instructional system. Construction of aptitude, diagnostic, survey, and pretests. These tests may be used to determine the knowledges, skills, and attitudes the prospective student has in his repertoire.

Plan, Develop, and Validate Instruction (Step 4): Involves sequencing learning activities in an order that produces the required learning in the shortest time. Selection of the instructional methods, media, and equipment. Development and validation of instructional materials. Validation indicates if instructional materials teach that which they are designed to teach and insures that all elements of the instructional system function effectively.

Conduct and Evaluate Instruction (Step 5): Determination of student attainment of learning objectives. This often identifies problem areas and the corrective action needed. Insures that the product of the instruction system (i.e., the qualified personnel) meet the requirements of the operating commands.

Figure 2. Definitions of the five Instructional System Development steps.

BEST COPY AVAILABLE

Regardless of the content, each interview first acquired routine demographic information. This information involved the interviewee's name and current assignment, his time in his current assignment, and an ATC course with which the interviewee was familiar and that had been developed in accordance with the ISD principles. The reason for the last item was that the interviewee was asked later about techniques and experiences in relationship to a specific course with which he was familiar. If the interviewee was not familiar with a course that was developed in accordance with the ISD model, then the course with which he was most familiar was identified and referred to during the interview.

The additional content of each of the four ISD interviews is described categorically below:

ISD Specialists

- familiarity and depth of experience with ISD
- the amount of emphasis placed on each ISD step in developing the course and if the step was not emphasized, why
- the success of certain ISD steps in developing the course
- how certain steps were helpful in developing the course
- identification of a situation where ISD steps were used but were not helpful in developing the course

Training Evaluation Specialists

- the frequency and method of feedback for course revisions from course review and field evaluation techniques
- evaluation of course review and field evaluation techniques
- problems with the course review and field evaluation techniques
- course review and field evaluation techniques that could be used to identify areas of over and undertraining
- identification of situations in which course review and field evaluation techniques were used to modify the course

BEST COPY AVAILABLE

Student Measurement Specialists

- types of student achievement measurement techniques used
- end of block written tests--evaluation, problems, determination of failing grades, recommendations to failing students, type of item analysis conducted, relationship to other measures
- performance tests--evaluation, problems, determination of failing grades, recommendations to failing students
- criterion checklists--evaluation, problems, determination of failing grades, recommendations to failing students
- other student measurement techniques that could be used
- feedback of student measurement results to course developers and how they were used
- identification of a situation where student measurement results were used to modify the course
- the frequency with which student measurements were reviewed and the frequency with which they were modified
- identification of a situation where student critiques were used to modify the course

Training Managers/Instructors

- the frequency and method of feedback for course revisions from course review and field evaluation techniques
- evaluation of course review and field evaluation techniques
- problems with course review and field evaluation techniques
- course review and field evaluation techniques that could be used to identify areas of over and under-training
- identification of situations in which course review and field evaluation techniques were used to modify the course
- types of student achievement measurement techniques used
- end of block written tests--evaluation, problems, determination of failing grades, recommendations to failing students

BEST COPY AVAILABLE

- feedback of student measurements to course developers and how they were used
- identification of a situation in which student measurement results were used to modify the course
- identification of a situation where student critiques were used to modify the course
- criteria for judging instructional effectiveness
- identification of the best and worst aspects of ISD and ranking them
- reasons for supporting the ISD program
- problems in developing courses using the ISD techniques
- evaluation of ISD techniques

In addition to the interview content, as outlined above, there were a series of items which pertained specifically to the ISD program and were included in all interviews. These items involved the following concepts:

- evaluation of the ISD program
- suggestions for improving the use of ISD
- reasons for avoiding ISD
- problems in using ISD
- evaluation of ISD steps for achieving an effective instructional program
- evaluation of ISD steps for the amount of emphasis given to each in developing the course

The final question in every interview, "Is there anything else you would like to tell me?" gave the interviewee the opportunity to free associate and in some cases to discuss aspects of training and ISD that may not have come out in the earlier parts of the interview. Some interviewees took the opportunity to clarify and elaborate on points that were made during the interview, and others discussed related subjects of interest and concern.

Reliability of Interview

Reliability, in terms of consistency of response, was investigated through inclusion of parallel items within each interview, i.e., items shich queried into the same issue but were stated differently

BEST COPY AVAILABLE

than their counterparts. Each of the five interview forms contained at least two pairs of items that were analyzed for response reliability. Examples of the reliability items are: What specific or general suggestions do you have for improving the use and application of ISD? and What one improvement would help ISD utilization most?; In general, would you rate or describe the ISD program as "superior," "very good," "good," "fair," or "poor" for achieving the stated objectives of turning out graduates who meet job performance requirements efficiently and effectively? and How would you evaluate the ISD program, on the basis of the values 0 to 100, for achieving its purposes?

Interviewer Training

Three interviewers collected the data in the field. The interviewers were all trained in psychology and the principles and techniques of structured and unstructured interviewing. Two interviewers were from the staff of Applied Psychological Services; the third was the project monitor. Training sessions were conducted with the interviewers. In these sessions, the purposes of the study, the areas of concern and interest, and the interview were discussed. All instructions to the interviewer, as they occurred in the interview, were reviewed, along with the purpose and meaning of each item. The manner in which the responses were to be recorded on the response sheets was covered. Other topics of discussion during the interviewer training sessions were: appropriate times to probe, methods for probing, how to handle reluctant interviewees, and organization and management of Air Force training courses.

Also included within the training sessions were considerations of typical problems inherent in the semistructured interview such as: interviewer bias, wording and posing questions, assuring the respondent that his responses would be used for research purposes only and would never be attributed to him, and methods for conducting a postinterview debriefing with each respondent.

Additionally, all interviewees familiarized themselves thoroughly with the reference documents which are listed on page 16 of this report.

Sample

The sample of interviewees included personnel from five different organizational or specialty areas. The sources from which the sample was drawn were: Instructional System Development, Training Manager/ Instructor, Student Measurement, Training Evaluation, and Headquarters.

BEST COPY AVAILABLE

Six primary sampling units were involved. These included five Air Force Technical Training Schools and Headquarters, Air Training Command (ATC) at Randolph Air Force Base, Texas. The five technical training schools are located at the Lackland, Sheppard, Keesler, Chanute, and Lowry Air Force Bases. The various specialists enumerated above were available at each of the five technical training centers, while headquarters personnel were available at Headquarters, Air Training Command.

Within specialties and centers, a random and proportional sampling approach was followed. This assured representation within the final sample of each specialty in proportion to its relative representation at each technical training center. Thus, each stratum represented in the total sample reflects its relative proportion in the population at each Air Training Command site. Table 1 presents, by job specialty and by Air Force Technical Training Center, the number of interviewees in the final sample.

Table 1

Sample of Respondents

Technical Training Schools

Job Specialty	Lackland	Sheppard	Keesler	Chanute	Lowry	Randolph	Total
ISD	10	6	7	7	3	-	33
TM/I	3	15	13	11	10	-	52
SM	7	4	5	9	5	-	30
TE	6	4	4	4	3	-	21
HQ ATC	-	-	-	-	-	3	3
Total	26	29	29	31	21	3	139

BEST COPY AVAILABLE

CHAPTER III

RESULTS

Description of Interviewees

The introductory aspect of each interview inquired into the background and experience of each interviewee. The initial item inquired into the length of time the interviewees were in their then current position. Figure 3 represents the mean percentage of each subsample by years in the current assignment for the four training specialties sampled. The fifth job specialty, headquarters personnel, was not included in the analysis because the headquarters sample consisted of only three interviewees. Figure 3 indicates that for each of the four specialties, the majority of respondents had at least from one to three years of experience in their current assignment. A very low percentage of respondents had less than one year of experience. The range of this percentage of respondents with less than one year on their current job was from 5.9 per cent to 15.6 per cent. The ISD job specialty had the largest percentage of respondents with less than one year on the job (i.e., 15.6 per cent). Since ISD is a relatively new specialty, it might be expected that a greater percentage of these specialists would possess minimum on-the-job experience.

Medians for the number of years in current assignment were also calculated. These were: SM* --2.4 years, TE* --4.2 years, TM/I* --2.4 years, and ISD* --2.1 years. With the exception of the TE group (which had a median of 4.2 years), the median values for the other three groups are very similar, with 2.1 to 2.4 years on the job representing the range of medians.

Since the interviewees were questioned about their opinions of the ISD program, an indication of their direct experience with ISD was also sought in the introductory questions. Figure 4 presents the percentage of each specialty in the sample that was familiar with a course developed using the principles of the ISD model. All of the ISD specialist interviewees were familiar with a course that was developed under the ISD program. Approximately three-fourths of the student measurement and training evaluation specialists were familiar with a course which was so developed. However, among the training managers/instructors sampled, only 41 per cent were familiar with such a course. In terms of the entire sample (across job specialties), 65 per cent were familiar with courses which had been developed in accordance with the ISD model. This overall percentage seems low in view of the advocated widespread use of the ISD

*SM-Student Measurement, TE-Training Evaluation, TM/I-Training Managers/Instructors, and ISD-Instructional System Development.

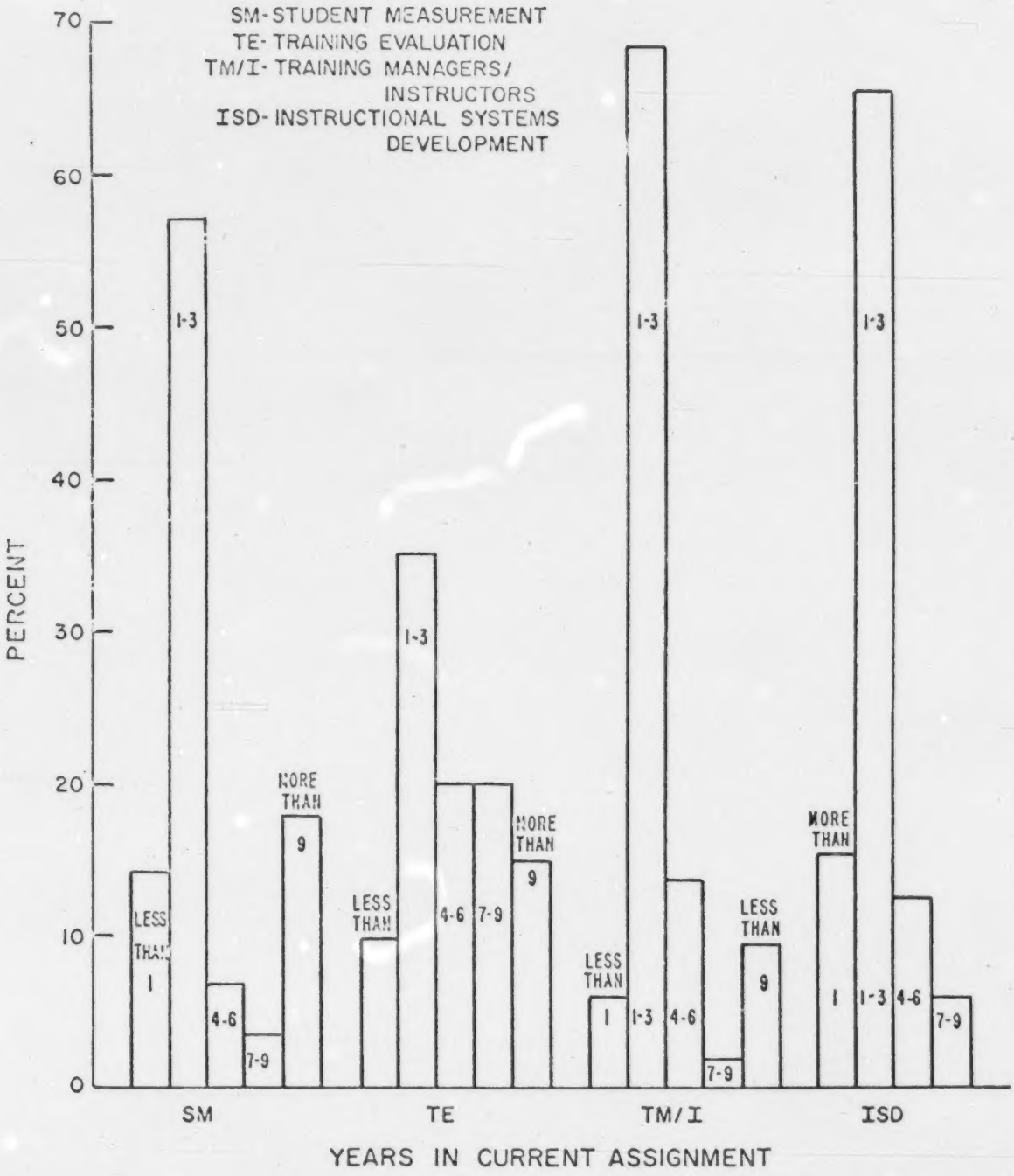


Figure 3. Mean time in current assignment for sampled interviewees in four training specialties.

BEST COPY AVAILABLE

system within the Air Force and suggests that many courses or course elements have been developed on some other foundation.

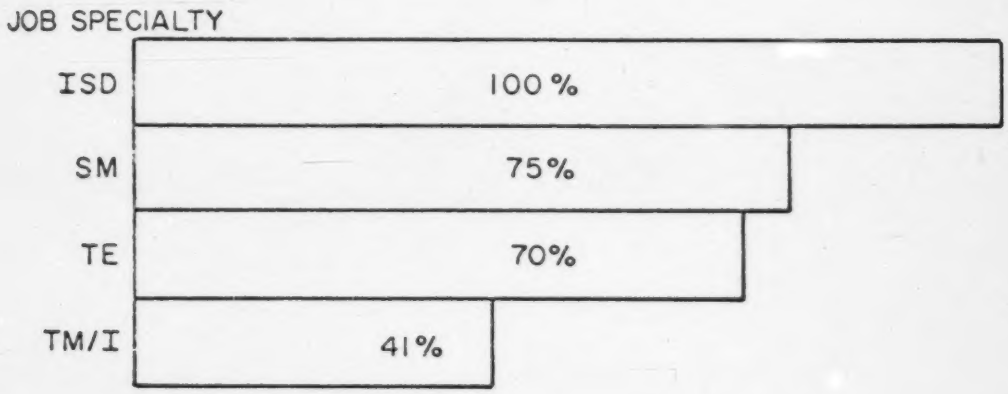


Figure 4. Percentage of each subsample possessing familiarity with courses developed in accordance with ISD principles.

Instructional System Development Specialists

Thirty-three ISD specialists were interviewed at the five Air Force technical training schools. The ISD interview consisted of two main sections. The first section of the interview dealt with the specific application of the ISD steps. The first substantive area that was investigated in this section of the interview was the amount of emphasis that was placed on each of the five ISD steps in developing the course with which the interviewee was most familiar and which had been developed in accordance with the ISD model. Then, the degree of success of two (the first and last) ISD steps for achieving their purpose was examined. The amount of emphasis and the degree of success were quantified through employment of the magnitude estimation technique (Torgerson, 1958). To this end, scales ranging from 0 to 100 were presented to the interviewees with the instruction that they select a value representing the amount of emphasis on each of the five steps and the degree of success for each of two steps.

BEST COPY AVAILABLE

Emphasis on Each ISD Step

Items 4, 11, 14, 17, and 20 of the ISD interview (Appendix A) pertained to the emphasis placed on each ISD step during the development of a course which had been developed through the ISD system approach. The means and standard deviations of the reported emphasis placed on each ISD step are presented as Table 2.

Table 2

Amount of Emphasis Placed on ISD Steps in Course Development

ISD Step	Mean	Standard Deviation
1. Analyze System Requirements	69	26
2. Define Education and Training Requirements	71	24
3. Develop Objectives and Tests	89	14
4. Plan, Develop, and Validate Instruction	80	21
5. Conduct and Evaluate Instruction	77	27

Table 2 indicates that, according to the scale of emphasis, the last three ISD steps tended to receive greater emphasis than the first two. Steps 1 and 2 had mean emphasis values of 69 and 71, respectively, while the mean values assigned to the remaining steps ranged from 77 to 89. The standard deviations for all but the third step (develop objectives and tests) seem sizable. For step 1 (analyze system requirements) those respondents who selected low emphasis values did so because the analysis of system requirements was performed outside the training department. In these cases, the course developers used either occupational specialty reports, expert opinions, and judgments, or data that were used for previous courses. On the other hand, those respondents who selected high emphasis values claimed: "step 1 was basic to the ISD program and therefore very important," "required great effort to complete," and "took a long time to do."

The respondents who placed little emphasis on step 2 (define education and training requirements) did so because they had either insufficient time, used job documents for defining the education and training requirements, or did not have to perform the details of the step because they were accomplished by other departments.

The interviewees who indicated that extensive emphasis was placed on step 3 (develop objectives and tests) offered reasons very similar to those offered for step 1. Some interviewees claimed that most of step 3 was accomplished while doing step 2 and 1 or that some new

courses were revamps of existing courses. In these cases, step 3 did not require too much emphasis. A few respondents indicated that the step 3 had been performed to satisfy regulations and/or requests made by their superiors. This suggests that, in these cases, step 3 was emphasized more than it would have been if the respondents had not been so directed. The most common reasons for step 3 receiving an extensive amount of emphasis was: "it is important," and "required if there is to be transfer to the field situation."

There were no respondents who claimed that very little emphasis was placed on step 4 (plan, develop, and validate instruction). Five respondents claimed that a moderate amount of emphasis was placed on step 4, whereas the rest of the respondents indicated that an extensive amount of emphasis was placed on this step. Respondents who placed a moderate degree of emphasis on this step indicated reasons such as: "the step is incomplete," "used material from previous course," and "not enough time allowed." The responses from the high end of the emphasis scale were, again, similar to those responses offered for the other ISD steps.

In regard to step 5 (conduct and evaluate instruction) some of the ISD specialists indicated that this activity was actually the responsibility of instructors and other sections. Accordingly, they placed relatively little emphasis on the step. Others indicated that the fifth step of the ISD model required much emphasis since it "is the crux of a good training program," and "since ISD is new, we had to be extra cautious." Other responses at the high end of the scale indicated that student critiques and field evaluations were constantly being reviewed for purposes of evaluating instruction, and these required a considerable amount of consideration and effort.

Finally, some of the ISD specialists replied that since they were responding in terms of the ISD based course with which they were most familiar and since the course involved was currently under development, there was little basis for their estimate in regard to step 5. This may account for the somewhat elevated standard deviation for this step.

Success of Each ISD Step

As was stated earlier in this section, the interviewees were asked to provide a magnitude estimate of the success achieved in implementing ISD steps 1 and 5. No inquiries into the success of steps 2, 3, and 4 were made because the interview pretest results indicated that the respondents assigned a very similar success value to each step. It seems that, to some extent, an overall halo effect was functioning in the pretest evaluations. This tendency was also inherent in the field survey, where the same mean success value (and very similar standard deviations) was obtained for both ISD steps investigated (Table 3).

Table 3

Amount of Success of ISD Steps 1 and 5

ISD Steps	Mean	Standard Deviation
1. Analyze System Requirements	84	17
5. Conduct and Evaluate Instruction	84	15

The mean value of 84 indicates that both steps 1 and 5 were considered to be extremely successful for implementing and controlling course effectiveness. Some of the reasons why the respondents considered step 1 to be successful are contained in the following typical responses: "helped to remove unnecessary material from the course," "an evaluation of the entire course indicated that the course was successful and therefore step 1 had to be successful," and "it brought about a closer alignment with the needs of the field." Step 5 was considered to be successful because: "it resulted in improved instruction," "it reduced course time making the course more cost-effective," and "by implication since the course worked well step 5 had to be successful."

Why Steps 1 and 5 Were Helpful

Item 8 of the interview asked how the first ISD step, analyze system requirements, was helpful in planning the course. Generally, the interviewees attributed most of the benefits derived from this step to result from the hierarchical ranking the analysis yields. This allows, according to the interviewees, more precisely defined course content, better identification of areas of emphasis, and courses which are structured so as to meet the job requirements.

The same question was asked in regard to step 5, conduct and evaluate instruction. Here, the interviewees indicated that this step had identified changes needed in the lesson plans, changes needed in time allocations, and changes in training methods or emphasis.

In no case could a respondent recall a situation in which steps 1 or 5 were used in developing a course and the course was not benefitted. Items 10 and 26 of the interview inquired specifically into this point, but were unproductive in that "no response" was recorded for all interviewees. This suggests that the application of the ISD principles to course development has met with considerable success wherever applied, as seen by over 30 ISD specialists assigned to the five Air Force technical training schools sampled.

BEST COPY AVAILABLE

General Opinions

The second main section of the interview with the ISD specialists was concerned with general opinions toward the ISD model, its use and application. Initially, the interviewees were asked to evaluate the ISD program for achieving the stated objectives of graduating students who meet job performance requirements efficiently and effectively. The results to the ISD evaluation item (item 27) are presented in Table 4.

Table 4

Results of Evaluation of the ISD Program by ISD Specialists (N = 33)

Classification	Per cent of Respondents
Superior	61
Very good	30
Good	6
Fair	0
Poor	0
No response	3

Evidently a large majority of the ISD specialists thought very highly of the ISD program. Sixty-one per cent claimed the program was "superior," in terms of achieving its stated objectives, 91 per cent of the responses were either "superior" or "very good."

In a parallel question, each respondent was asked to evaluate the ISD program along a numerical scale which ranged from 0 to 100. The evaluation was in terms of the extent to which the program achieves its purpose. The mean of this evaluation was 78, with a standard deviation of 20. This mean value is highly concordant with the categorical replies discussed above. (If one assumes each of the five categories, discussed above, to represent a 20 point range, then the expected value for the magnitude estimate is [(61 x 90) + (30 x 70) + (6 x 50)/100 = 78.90)]. The close accord between the responses to these two questions, one which was asked at the start of the interview and one which was asked near the end, supports a contention favoring the reliability of the interview responses.

BEST COPY AVAILABLE

Problems Inherent in ISD Technique

The next area of investigation pertained to the problems that were met by the respondents in using the ISD model and suggestions that the interviewees thought might be employed to improve the application and utilization of ISD. The most frequently mentioned problem was concerned with the lack of adequately trained personnel. This implies both a need for more personnel and for training these persons in ISD methods and techniques.

A second frequently mentioned problem was insufficient time to complete the various ISD steps. This problem arises as a result of several factors. For some course developers, it is very possible that unrealistic deadlines (for the completion of individual ISD steps) were set so that the most proficient course developer could not finish in the time allowed. Other course developers may not have been able to meet deadlines due to a lack of training in ISD. Still other course developers may not have had sufficient personnel or technical support available.

A third set of problems was concerned with a variety of characteristics which might be anticipated in any evolving program. For example, several respondents pointed to insufficient funds as the source of their difficulty with the ISD program. Others mentioned inadequate equipment (e.g., training aids/devices, training facilities and documents).

These problems are real to the ISD specialist and, in some cases, can eclipse the many positive attributes of the ISD program. Accordingly, solutions to these problems seem required if the ISD program is to mature fully. This implies the development of a more thorough and complete training program in the use and application of ISD, an increase in the number of trained ISD specialists, and increased information dissemination regarding the ISD program.

Other suggestions offered by the respondents included: (1) re-performing or performing steps 1 and 2 for existing courses, (2) developing screening and selection criteria for ISD specialists, and (3) improving the ISD documents so that they are more readable.

A related question asked why people avoid using the ISD method. Here, the respondents seemed to think that: (1) many course developers and especially the older ones are more comfortable and secure performing in their accustomed manner and they, accordingly, resist change, (2) there is too much ignorance regarding ISD, and (3) course development through the use of the model requires too much time.

BEST COPY AVAILABLE

Training Evaluation Specialists

Twenty-one training evaluators were interviewed for the purpose of obtaining information in two principal areas: (1) ISD methods and procedures, and (2) training evaluation.

ISD Evaluation

The training evaluation specialists were asked to evaluate the global ISD program on a categorical scale which ranged from "poor" to "superior." This evaluation was made, as for the ISD specialists, in terms of how well the ISD method achieved its goal of graduating students who meet job performance requirements efficiently and effectively. Table 5 presents the result of this evaluation.

Table 5

Results of Evaluation of the ISD Program by Training Evaluation Specialists (N = 21)

Classification	Per cent of Respondents
Superior	10
Very good	43
Good	29
Fair	14
Poor	0
No response	4

A small percentage (i.e., 10 per cent) of the training evaluators rated the ISD program as "superior" and 43 per cent rated it as "very good." The remaining responses fell in the "good" and the "fair" categories. Here, there is a lower percentage of respondents employing the two higher categories than for the ISD specialists. This trend may have been anticipated since the ISD specialists, in view of their more total personal involvement in the first four steps of the ISD sequence, would be expected to provide a more favorable self-evaluation.

When the training evaluation specialists rated the ISD program (for achieving its purposes) on the quantitative scale, ranging from 0 to 100 (item 22), the obtained mean rating was 58, with a standard deviation of 19. This value is similarly lower than the corresponding value (78) yielded by the ISD specialists. (The anticipated magnitude estimation value, for the training evaluators on the basis of the replies to the categorical question, was $[(10 \times 90) + (43 \times 70) + (29 \times 50) + (14 \times 30)/100 = 55.9)]$.

BEST COPY AVAILABLE

Problems Inherent in ISD Technique

The training evaluation specialists also most frequently mentioned the lack of adequately trainer personnel as the major problem within the ISD technique. The other frequently mentioned problems with ISD were that: (1) it is extensively time consuming, (2) occupational survey data are not readily available (i.e., the job analytic data required in step 1 of the model) for updating a course, and (3) there is a lack of coordination between the ISD and training evaluation efforts with respect to course evaluations.

Inquiry into solutions for these problems resulted in suggestions such as increasing the accessibility of the ISD course to ISD specialists so as to increase the number of trained specialists at the training centers, and the development of courses in the concepts and benefits of ISD for personnel associated with the Air Training Command, but not assigned to the ISD specialty. It was also suggested that streamlined procedures be developed for course development under the ISD principles. The streamlined procedures would result in a decrease in the time needed to prepare a course. The total time needed for ISDing a course could also be reduced, as the respondents suggested, by increasing the manpower assigned to developing the course.

Other considerations for improving ISD pertained to the analyses of system requirements that are performed in the first step of the model and an information dissemination program which would help to positively motivate people in managerial positions in regard to the ISD program. The former consideration, that of developing a job analysis, is crucial to successful course development under the ISD precepts. In some cases, the job analyses were not available, and existing data and/or ad hoc analyses were performed at the training centers or in conferences, rather than in the field. The latter consideration, information dissemination and diffusion, was not offered in the sense of propagandizing or attributing qualities to ISD that are not factual. Rather, the intent was to inform and educate people about the ISD methods, how they are used, and the inherent advantages of the systems approach to training program development.

The reason people have avoided using ISD, from the training evaluator's point of view, was the ever present attitude of resisting change when the need for change has not been aroused, a lack of knowledge and understanding of the ISD system, and because full application of the ISD methods is time consuming and somewhat costly.

Training Evaluation

A definition of training evaluation was read to the training evaluation specialist respondents at the outset of the interview.

BEST COPY AVAILABLE

Training evaluation was defined "...as the determination of how well learning objectives were identified, the efficiency of the ISD process, and the proficiency of graduates in performing the jobs for which they were trained." A set of questions was then presented to the interviewees, relative to training evaluation.

The first four items concerned the updating of courses to meet changing requirements. All respondents, but two, were involved with courses that had been periodically changed or modified. Most respondents indicated that the courses were modified, on the average, about once a year. However, it appears as though some courses were modified as frequently as at the end of each course, and as infrequently as every other year. A percentage breakdown of the responses indicated that in approximately 50 per cent of the cases courses were updated about once a year; in 20 percent of the cases there were revisions at the end of each course, and in another 20 percent courses were updated semi-annually; 5 per cent underwent revisions once every other year and the remaining 5 per cent could not be pinpointed. Course developers were usually informed that a course revision was needed as the result of information provided in student critiques, written reports (e.g., field evaluations, periodic course reviews), and conferences.

Course Review Techniques

The typical course review techniques, as separate from field evaluation, used were student critiques/student interviews, supervisory reviews, and reviews of field requirements. Less typical course review techniques mentioned were instructor reviews and course document reviews.

The respondents were also asked to evaluate the course review techniques used for obtaining information that could be used for modifying a course. The results to this evaluation are presented in Table 6.

Table 6

Evaluation of the Course Review Techniques Used
(N = 21)

Classification	Per Cent of Respondents
Superior	5
Very good	32
Good	29
Fair	29
Poor	5

BEST COPY AVAILABLE

A small percentage of the training evaluation respondents regarded the techniques used in the courses with which they were most familiar as either "superior" or "poor." The majority of the responses tended more toward the middle of the scale. Approximately one-third of the interviewees indicated that the techniques used in course reviews were "very good." An equal percentage felt the techniques were "good" and again, about one-third thought they were "fair."

Insights into the rationales behind this less than optimum indication were sought through the question, "What was the principal problem with the course review technique employed?" The single most frequently identified problem was the subjectivity or biases that were allowed to creep into the evaluations. A close followup to the problem of subjectivity was that some course reviewers were not skilled in the course review techniques employed. These comments suggest the need for training programs which would prepare course reviewers in the techniques of course evaluation including the use of objective techniques (such as rating scales, checklists, etc.). Some Techniques for the Evaluation of Technical Training Courses and Students (AFHRL-TR-72-15), developed by Applied Psychological Services as a separate part of the current program, may also help to provide at least a partial solution in this regard.

Other problems mentioned as inherent in the course review techniques used at the training centers were: cumbersomeness of certain training evaluation forms and, in some cases, performance of course reviews by the technical school personnel rather than by the training evaluators. The solutions to these problems are self-evident.

Identification of Areas of Over and Undertraining

The problem of over and undertraining is of concern to any training program that has cost effectiveness as a goal. The training evaluators were, accordingly, asked which course review technique, as separate from field evaluations, best identified areas of over and undertraining. The training evaluators generally thought that while in-house course review techniques might identify areas of undertraining, no course review technique, as such, could be used for identifying areas of overtraining. Rather, they suggested that an experimental program associated with achievement tests be undertaken to determine areas of successful performance. For example, if as a result of an achievement test, it was discovered that the students were all performing satisfactorily in a particular subject matter area, then perhaps this area could also be performed satisfactorily if slightly less time were spent on it during the course. The time saved could be devoted to other subject matter that the students find more difficult to master, or the course could be commensurably shortened.

BEST COPY AVAILABLE

The training evaluators also indicated that the subjective comments contained in student critiques and instructor suggestions often identify areas of over (as well as under) training.

Specific Use of Course Reviews

The interviewees were further asked if they were aware of any specific incident in which a course review identified a problem and in which corrective action was taken. The respondent described the problem and stated how it was handled, who was responsible for instituting the correction, and how the modification was beneficial. Approximately 30 per cent of the interviewees responded to this item. In about three-fourths of the incidents related, the Training Evaluation Divisions of the respective technical training schools were responsible for introducing a change in either the course content or course administration. The other one-fourth of the incidents were handled by instructors and instructor supervisors and, as before, either course content or course administration was affected. For example, in one incident the training evaluation department found certain lectures of a course to be dull and uninteresting. Students were interviewed, and this finding was verified. The training evaluators suggested to the instructors that the presentation of the subject be made more crisp and interesting. The instructors decided to employ the programmed instructional technique in place of the routine lectures. In addition to an increased interest level, a time savings was introduced.

Field Evaluation

Field evaluation, as separate from course reviews, was inquired into through a separate set of questions. The field evaluations usually involved critiques offered by both graduates of the course and the supervisors of these graduates. Graduates usually critiques the training received at the school, whereas the supervisors critiques the graduates' ability to perform on the job. The respondents indicated a program of posttraining performance testing to be a field evaluation technique in only rare instances (e.g., three respondents indicated that posttraining testing was used as a field evaluation technique in the course on which they were reporting).

The results of the field evaluations usually were fed back to the training evaluators through written reports or by way of the questionnaires that course graduates and supervisors completed. Another way by which training evaluators were made aware of field evaluation results was through meetings.

The adequacy of field evaluations, as a technique for obtaining information that can be used for modifying courses, was evaluated by the training evaluators. Table 7 presents the result of this evaluation.

BEST COPY AVAILABLE

Table 7

Evaluation of the Field Evaluation Techniques Used
(N = 21)

Classification	Per Cent of Respondents
Superior	14
Very good	48
Good	33
Fair	0
Poor	0
No response	5

Table 7 indicates that field evaluation techniques were never placed in the lowest categories, "fair" or "poor." The interviewees were obviously favorably disposed toward the field evaluation techniques, with 62 per cent claiming they were "very good" or "superior" and 33 per cent indicating that they were "good."

There is a striking difference between this evaluation of the course review techniques and evaluation of the field evaluation techniques (cf. Tables 6 and 7). Whereas none of the respondents thought of the field evaluations as "fair" or "poor," 34 per cent thought the course review techniques were classified as either "fair" or "poor." Sixty-two per cent of the respondents classified the field evaluation techniques as "superior" or "very good," but only 37 per cent thought similarly of the course review techniques.

When the interviewees were asked to state the principal problem with the field evaluation techniques, the "careless" and "indifferent" manner in which the questionnaires are completed was most frequently mentioned. The other problems mentioned were that small samples were used and that the comments were very subjective. The respondents, in many cases, claimed the superiority of the personal interview over the mail questionnaire because interviews allow fuller exploration of the basis for remarks. Personal interviews are, however, more time consuming and demand a skilled interviewer. Accordingly, the questionnaire technique is often adopted as a matter of expediency. However, it is possible that modification of presently employed questionnaires might produce more meaningful information.

The interviewees indicated that evaluations produced by the graduates yielded the most meaningful and useful data--both in regard to areas of overtraining and areas of undertraining. This does not imply that supervisory evaluations were not regarded as having value in this area, but rather that they provided less meaningful information than that provided by the graduates.

BEST COPY AVAILABLE

Approximately 60 per cent of the sample responded to the critical incident type item (item 17) which inquired into incidents in which field evaluation results were used by course developers. In all but three cases, the training evaluators were responsible for instituting the resultant change; instructors and headquarters personnel were responsible for modifying the course in the three cases. In every incident related, the course content was modified, and the modification was found reasonable and profitable. An example of a specific incident follows. As a result of a field evaluation, it was discovered that graduates of a course were not trained on one of the two computers used in the field. Training evaluators suggested that the course be updated to cover both computers. The instructors modified the course to include the computer that had been omitted.

Student Measurement Specialists

Thirty student measurement specialists were sampled at the five Air Force technical training schools visited. The interview with the student measurement specialists consisted of two major sections. One section dealt with the specifics of the student measurement techniques employed at the training centers; the other involved an evaluation of the ISD program. The ISD section of the interview will be discussed first.

ISD Evaluation

The student measurement specialists regarded the ISD program with moderate favorableness for achieving the objective of graduating students who meet job performance requirements efficiently and effectively. When asked to rate the ISD program on the five category scale which ranged from "superior" to "poor," 10 per cent of the student measurement specialists placed the program in the bottom two categories, "fair" and "poor." Forty-seven per cent considered the ISD program as "superior" or "very good," and 33 per cent considered the program as "good." The results to this evaluation are presented in Table 8.

BEST COPY AVAILABLE

Table 8

Results of Evaluation of the ISD Program
by Student Measurement Specialists (N = 30)

Classification	Per Cent of Respondents
Superior	20
Very good	27
Good	33
Fair	7
Poor	3
No response	10

When later in the interview the respondents were asked to evaluate the ISD program on the quantitative scale which ranged from 0 to 100, the mean obtained value was 62 with a standard deviation of 22. This value may be compared, for reliability purposes, with an expected value (on the basis of the categorical data) of 55.8. This value (56) is again lower than that assigned by the ISD specialists (78) and is identical to that assigned by the training evaluation specialists.

The problems most frequently mentioned by the student measurement specialists in regard to the ISD model were: insufficient number of trained ISD specialists, inadequate general information program on the advantages of the ISD technique, and difficulty with certain of the ISD steps.

Again step 1 (analyze system requirements) was said to cause difficulty, and in some cases task analyses had not been performed for courses that were otherwise modified in accordance with the ISD model.

The student measurement specialists suggested several solutions to the problems that were experienced in applying the ISD method. Most frequently mentioned for improving the use and application of ISD was, again, the development of more and better trained ISD specialists. Other suggestions for improving the use and application of ISD were: (1) increase the flow of information in regard to the ISD program so that the training personnel responsible for developing courses on the basis of the ISD method would meet with less opposition and with greater cooperation, (2) allow more time to prepare a course when using the ISD method, and (3) reduce the volume of required paperwork.

BEST COPY AVAILABLE

The need for an increase in the time allowed to prepare a course by using the ISD may be a result of the fact that the technique has often been new to the personnel assigned to course development. As ISD specialists and others assigned to developing courses gain experience with the ISD method, it is likely that less time will be required to develop courses with the procedure.

There seemed to be agreement among the student measurement specialists on the reason that the ISD method may sometimes be avoided. The student measurement specialists, very much like the other specialty groups, thought the ISD method should be avoided because course development personnel were resistant to change and did not understand the method. Moreover, they claimed that use of the ISD method involves too much work. Some interviewees were of the opinion that if more ISD specialists were available, the amount of work required from any one person in developing a course would be diminished, commensurate with the increased number of ISD specialists.

Student Measurement

Prior to questioning the student measurement specialists in the area of student measurement techniques, a definition of student measurement was read. The definition stated, "We define student measurement as measures for the purpose of evaluating individual student achievement and progress." Table 9 presents the results to the first question in this area. This question asked the interviewees to identify the different student measurement techniques that were employed in the course they identified at the beginning of the interview as being the one with which they were most familiar.

Table 9

Student Measurement Techniques Employed
(N = 30)

Technique	Number of Responses
End of block written test	28
Performance test	26
Criterion checklist	18
Instructor rating	13
Programmed instruction test	7
Oral test	7

BEST COPY AVAILABLE

Table 9 indicates that the end of the block written test and the performance test are the most frequently employed student achievement measurement techniques employed. The criterion checklist, which is gaining in popularity, was identified as a technique used by somewhat fewer student measurement specialists. Instructor ratings were used, but less frequently than the prior three techniques, and the programmed instruction test and the oral test were relatively infrequently used.

End of Block Written Tests

A set of subsequent items inquired specifically into end of block written tests. Absolute scores (as opposed to criterion referencing or norm referencing) were used to determine the passing point in all cases. Item analyses, for the purpose of appraising the usefulness and effectiveness of each item, were conducted in all courses where end of block written tests were used. They were typical item analyses in which difficulty indexes are obtained. In only a few cases, the respondents indicated, were the end of block written test results correlated with other measures of performance to determine the degree of relationship between two different measures of performance. In the few cases where these correlations were obtained, the end of block written tests were correlated against performance test scores.

Although many respondents thought that end of block written tests were effective student evaluation devices, many also thought that they were ineffective tools. A small number of the respondents reported that end of block written tests had motivational value (i.e., students faced with written tests will study for them). These tests were also held to be comprehensive, practical, and objective (i.e., the short answer type, rather than the essay type).

In this set of questions, the respondents were also asked to identify problems involved in the use and development of these tests. The major problem identified was test construction difficulty (e.g., writing effective items). Students were always informed of their end of block written test results. Those who did not achieve satisfactory results usually were required to repeat the block. In some cases, special assistance was provided to the student who failed a written test. The student was then retested. If failure still resulted, the instructor could recommend that the student be dropped from the course.

Performance Tests

Performance tests were investigated in a manner similar to that in which the end of block written tests were investigated. Performance tests (sometimes referred to as work sample tests) were employed

BEST COPY AVAILABLE

in all but four of the courses involved. In some cases, absolute scores were used to separate passing and failing students, but in almost as many cases passing or failing was norm referenced. The respondents felt that the value of performance tests rested in their practicality and their use as a tool for keeping track of the progress of a student. Practical, in this context, has several meanings. It could mean that actual equipment, i.e., the same equipment that is used in the field was used in the test, or that the tasks and/or problems were closely related to those found in the field, or that if skills were taught in the course the performance test was a valid instrument for measuring achievement.

The most frequently mentioned problems with the performance tests were of a test administrative nature. Unlike a written test, the performance test allows only one student to be tested at a time. Thus, unless the classes are small or there are several test administrators, the performance test has to be of limited duration; otherwise testing student achievement can require considerable time. Three other performance test related problems were mentioned with some frequency: equipment for performance testing is sometimes difficult to acquire, the lag between equipment requisition and equipment receipt is too long, and performance test scoring is often subjective in nature. In regard to equipment acquisition problems, it seems that inexpensive mockups and/or paper-and-pencil performance tests might offer some solution.

As with the end of block written tests, students were informed of their performance test results. Those that did not receive satisfactory results received additional training or special assistance and/or were "washed back." Only in very few cases (where the performance was exceptionally poor) were failing students recommended for elimination from a course on the basis of the performance test results.

Criterion Checklists

The subsequent set of items involved the application of the criterion checklist. Two-thirds of the student measurement respondents were familiar with courses in which the criterion checklist was employed as a student measurement technique. In most cases, satisfactory and unsatisfactory performance was determined on a criterion referenced basis. However, some respondents reported that satisfactory-unsatisfactory performance was determined by subjective evaluation, and a small number of respondents replied that the criterion checklists they used had numerical scores given for various aspects of performance and an absolute score was used to determine a student's final score.

BEST COPY AVAILABLE

The criterion checklist was almost universally said to represent a thorough, practical, and valuable technique for keeping track of student progress. In some courses, the criterion checklist was used as a precursor to the end of block written test. Here, students had to perform satisfactorily the items on the criterion checklist before he could take the written test. Thus, the checklist became a teaching aid as well as an achievement measuring instrument. As with the performance test, the criterion checklist was seen as practical because the students were called on to perform as they would in the real situation. Also, the student was evaluated by an instructor who, in this situation, could function in much the way that a recent graduate's supervisor would function.

The problems inherent in the criterion checklist fell into the following categories: administrative, scoring subjectivity, costly, and difficult to construct. For example, some interviewees claimed that they had to develop the criterion checklist before they had knowledge of the system requirements (step 1). This resulted in great difficulty in developing criterion objectives. Perhaps this problem is related to the sophistication of the test developer, his knowledge of the subject matter, and the needs of the field.

Some of the same problems inherent in performance test development and administration were also involved in the criterion checklist. In regard to test administrator reliability, some interviewees suggested supervision of the evaluator while he completes the checklist or obtaining evaluations from two independent evaluators. Other solutions to this problem might be sought through training programs in which problems of evaluator bias, approaches to evaluating performance, and completing the checklist are fully explored.

As with end of block written tests and performance tests, students were always informed of their criterion checklist results. Those who did not achieve satisfactory results were either recycled or received additional training or special assistance. They then repeated the test they had failed. Only in rare circumstances did an interviewee report that a student, on the basis of criterion checklist results, was incapable of completing the course satisfactorily. But, if necessary, students were dropped from the course on the basis of criterion checklist results.

The retest procedure may suggest the need for parallel tests rather than repetition of the same test.

BEST COPY AVAILABLE

In summary, the perceived advantages and disadvantages of the three student achievement measurement techniques, as seen by student measurement specialists, are presented below:

Test	Advantages	Disadvantages
Written	1. keeps track of student progress 2. objective 3. motivating 4. thorough	1. difficult to construct 2. can be a test of reading ability
Performance	1. keeps track of student progress 2. practical	1. too long to complete if individually administered 2. subjective scoring 3. difficult to obtain requisite equipment
Criterion Checklist	1. keeps track of student progress 2. practical 3. thorough	1. too long to complete if individually administered by one test administrator 2. subjective scoring 3. difficult to construct

Needed Student Measurement Techniques

Item 29 of the interview with student measurement personnel queried into student measurement techniques which are not currently used but which the student measurement specialists would like to see used. The results indicated that the testing devices currently being used were considered to be satisfactory. This may suggest some lack of knowledge of more advanced testing practices such as confidence testing methods, computer assisted testing, hierarchical testing and the like.

Feedback to Course Developers

Test results were usually provided to course developers by way of student records, written reports of various kinds, school records, word of mouth, and in some instances at meetings. About 20 per cent of the interviewees reported instances in which the course developers had instituted corrective action on the basis of test results. The test results were used by the course developers in several different ways. If course changes or revisions were instituted as a result of student evaluations, the change was in the form of modifications to existing examinations or equally often in the form of changes in instructional techniques. On a less frequent basis, course outlines were revised or

BEST COPY AVAILABLE

learning objectives were modified. This last type of modification makes little sense to the present authors, if the learning objectives are derived from the performance requirements. On a very infrequent basis, student evaluations were used by course developers to modify the selection criteria for the course and to evaluate the instructor. This latter employment may also represent a misuse of test results because such use encourages the instructor to teach the test.

An example of a situation in which a modification was made as a result of text findings involves an item analysis of a written test. It was found that many of the high scoring students were missing a particular item on the test. The instructor and instructor supervisor decided to reword the item. The item demonstrated improved discriminability in subsequent test administrations.

Review and Revision of Student Measurement Techniques

When inquiry into the frequency with which student measurement techniques were reviewed and modified as a result of the reviews was made, 50 per cent of the respondents indicated that student measurement techniques were reviewed at the end of each course, 25 per cent claimed they were reviewed "aperiodically," 13 per cent said yearly, six per cent claimed that they were reviewed quarterly, and another six per cent said "semiannually." Evidently, firmer ground rules are needed here.

Modifications to measurement techniques, resulting from these reviews, occurred as frequently as at the end of each course for 17 per cent of the interviewees and semiannually for another 17 per cent. Twenty-three per cent of the respondents could not fix a frequency for the modifications that resulted from student measurement reviews, but indicated that modifications did result from the reviews. Seven per cent claimed that modifications to the tests were made about every quarter and another 7 per cent claimed that they were made yearly. Finally, 3 per cent of the interviewees replied that modifications were made about once every other year. The remaining 26 per cent of the interviews did not respond because they had insufficient information on this matter. Reviews of student measurement techniques can be profitable, to the degree that they yield modifications (if required) to the measurement devices. On the other hand, it seems that modification on the basis of the results from one class might represent an insufficient sample for firm conclusions. It would seem that the 17 per cent of the respondents who modify tests semiannually are performing this update at an acceptable frequency; others may be performing the update too frequently or not frequently enough.

BEST COPY AVAILABLE

Student Critiques

Student critiques were completed in all of the courses discussed by the interviewees. The critiques were reviewed by many different persons, sections and departments at the schools, from the instructors up to the base commanders. About one-third of the interviewees related an incident in which the course developers took corrective action on the basis of the critique forms completed by the students. In these incidents, the course supervisors were most often identified as being responsible for the corrective action. However, one incident was related in which the instructor was responsible for the change and two incidents involved other departments at the school. The incidents were varied so that the resulting corrections were applied to the course content, tests, and physical facilities. For example, in one instance, a student complained that the lighting in the laboratory was inadequate. The tasks performed in the laboratory required a high degree of precision, i.e., soldering connections. The lighting in the laboratory was measured and a work order requesting an increase in the lighting was submitted. The lighting level was increased, with the result that this unfavorable condition is no longer a negative factor affecting the learning environment.

Training Managers/Instructors

The interview for training managers/instructors consisted of three sets of items: ISD, training evaluation, and student measurement.

ISD Evaluation

Initially, the interviewees were requested to evaluate the ISD program for achieving its stated objectives of graduating students who meet job performance requirements effectively and efficiently. The results to the item are presented in Table 10.

BEST COPY AVAILABLE

Table 10

Results of Evaluation of ISD Program
by Training Managers/Instructors
(N = 52)

Classification	Per Cent of Respondents
Superior	10
Very good	44
Good	23
Fair	6
Poor	0
No response	17

Seventeen per cent of the respondents were realtively unfamiliar with the ISD program and therefore could not evaluate the program. Of the remaining 83 per cent, 54 per cent placed the program in the "very good"/"superior" classification, 23 per cent classified the program in the middle category "good," and 6 per cent considered the ISD program as "fair." None of the respondents considered the ISD method to be "poor."

As for the other groups interviewed, the general evaluation of ISD was also approached through the magnitude estimation technique in which the respondents were asked to evaluate ISD on a scale which ranged from 0 to 100. This evaluation yielded a mean value of 69 with a standard deviation of 18. Obviously the training managers/instructors regard the ISD favorably. (The anticipated mean value for the quantitative scale, on the basis of the categorical scale, is 53. Here, there is a greater difference between the anticipated and the obtained than for the other training specialties interviewed. This suggests greater response unreliability for this interviewee group than for the other groups). This value (69) places the reactions of the training managers/instructors to be lower than those of the ISD specialists but above those of the student measurement and the training evaluation specialists.

Typical problems (in order of frequency of response) reported by the training managers/instructors in developing new courses using the ISD Model included: difficulty in meeting the requirements of the ISD steps, inadequate understanding of ISD steps, lack of information, lack of personnel, insufficient time to prepare courses, inflexibility of system, and lack of cooperation from instructors.

BEST COPY AVAILABLE

The suggestions for improving the ISD program offered by the training managers/instructors were very much like those of the other specialties interviewed. Among the suggestions were: more and better trained ISD specialists, increased acceptance of the program at the training centers, more complete job analysis by going into the field instead of using mail questionnaires, assignment of subject matter specialists to work with ISD specialists (this will tend to speed up the process of preparing a course because it will relieve the ISD specialist from taking the time to learn the subject matter), improving the administrative aspects of ISD by reducing the amount of paperwork required, and raising the screening and selection criteria so that the Air Force will have a superior corps of ISD specialists.

The training managers/instructors maintained that ISD is avoided because it is time consuming, because its objectives are not understood, and because application personnel are resistant to change. Most training manager/instructors supported the employment of ISD for such reasons as: "it presents an efficient means of training people to meet job performance requirements," "less slipshod and therefore better courses are developed," and "the graduates are more appropriately trained." On the other hand, several interviewees indicated that their support for the ISD program was based solely on the fact that they were ordered to use it.

Best Aspects of ISD Method

During one aspect of the interview, the training manager/instructor was presented with an adjectival list of 10 words, all positively keyed, from which they were asked to select the three adjectives they thought best described the ISD procedures. Table 11 presents the list of adjectives, the number of times each adjective was selected, and the resultant hierarchical rank order.

Table 11

Positive Adjectives Selected to Describe the ISD Model

Adjectives	N	Rank Order
Organized	29	1
Effective	27	2
Efficient	14	3.5
Profitable	14	3.5
Complete	13	5
Appropriate	11	6.5
Helpful	11	6.5
Good	5	8
Dynamic	4	9

BEST COPY AVAILABLE

Table 11 indicates that the four most frequently selected adjectives were: "organized," "effective," "efficient," and "profitable." The adjectives appearing at the bottom of Table 11 were not necessarily thought of as inappropriate, but rather as less appropriate than those at the top of the list. Quite obviously, the possibility exists that the provided list was incomplete and that the respondents might have misunderstood the meaning of any one or more of the adjectives.

This item was followed by one which requested that the interviewees state, in their opinion, the three best aspects of the ISD program and to rank order the aspects from 1 to 3 (the rank order of 1 was the highest; 3 the lowest).

The responses were categorizable in six separate categories. The frequency with which each category received a ranking of 1, 2, or 3 was tabulated, and a total score (rank order) for each category was calculated. Table 12 presents these data.

Table 12

Rank Order of Best Aspects of the ISD Model

Category	Frequency*			Score
	1	2	3	
Systematic, standardized, organized, logical, validates instruction	17	16	13	96
More appropriately prepared graduates	10	5	6	46
Job oriented	8	7	3	41
Identifies course objectives	3	10	5	34
Improved tests and measurements	3	3	6	21
Cost effective	2	1	4	12

* Frequency of a ranking of 1, 2, or 3. The total N for each column varies as a function of the number of responses. Some interviewees identified only one or two of the ISD model's best aspects.

BEST COPY AVAILABLE

In order to obtain the scores (Table 12) a weighting scheme was used such that a rank order of 1 was weighted with a 3, the rank order 2 was given the weight of 2, and the rank order 3 was weighted with a 1. The weights were used as multipliers and the three products were added together to arrive at the final score for the category. Thus, for the first category listed in Table 12, the score of 96 was obtained as follows: 17(3) + 16(2) + 13(1) = 96.

The first category, the category with the highest score, pertains to the methodology inherent in the ISD model. Therefore, the training managers/instructors were most favorably impressed with the ISD methods. The least impressive aspect (but nevertheless among those identified as the "best") was cost/effectiveness. The results of this analysis are interesting from the point of view that the respondents, as training managers/instructors, should be concerned more with the methodology of a course development system than with some other aspect of the approach.

Worst Aspects of ISD Method

In a parallel manner, the interviewees were queried on the negative aspects of the ISD method. Initially, they were asked to select the three worst aspects of the ISD model from a list of ten adjectives. The results are presented in Table 13.

Table 13

Negative Adjectives Selected to Describe the ISD Model

Adjectives	N	Rank Order
Tedious	29	1
Costly	26	2.5
Limited	26	2.5
Static	6	4
Mundane	3	5
Unsuitable	2	7.5
Unworkable	2	7.5
Unsuccessful	2	7.5
Weak	2	7.5
Poor	1	10

BEST COPY AVAILABLE

Within the listed adjectives presented, the three worst aspects of ISD were that it is "tedious," "costly," and "limited." The other adjectives, although negative in connotation, were hardly selected by the respondents. The choices of "tedious" and "costly" are compatible with the responses of other groups to various aspects of the interview, as well as with the responses of the training managers/instructors to other questions. The choice of "limited" suggests that, to some extent, the respondents believed the ISD technique not applicable to all courses.

The training manager/instructor respondents were also asked to identify the three worst aspects of the ISD model, using their own words, and to rank order the aspects. The ranking of 1 was equal to the worst and 3 was equivalent to the least bad. The data were treated in the same manner as was used for best aspects, described above. The results are presented in Table 14.

Table 14

Rank Order of Worst Aspects of the ISD Model

Category	Frequency*			Score
	1	2	3	
Developing curricula, objectives, and tests	9	8	10	53
Costly and time consuming	9	6	8	47
Limited to teaching skills, not enough student-instructor interaction	10	5	2	42
Instructors and personnel inadequately trained in ISD methods	6	6	2	32
Analyzing system requirements	4	4	2	22
Feedback from field is slow	1	2	1	8

* The total N for each column varies as a function of the number of responses. Some interviewees identified only one or two of the ISD model's worst aspects.

BEST COPY AVAILABLE

As suggested in the responses, the most faulty category of ISD, according to training managers/instructors, was that skill and experience is needed to develop a course curriculum, objectives and tests by the ISD standards. Although the feedback that was built in to the model did not work to the satisfaction of the TM/Is, it was on the bottom of the list.

Evaluation of ISD Program

One interview question requested information on the procedures, if any, that were instituted for evaluating the ISD program. Most of the respondents were unaware of any such procedure. However, several indicated that certain formal procedures were used, i.e., cost comparisons and comparison's of graduate proficiency between ISD trained graduates and others. Evidently, increased emphasis is needed in this regard.

Course Review

The opinions of the training managers/instructors in the course review sphere were sought in a separate set of questions. All but three of the interviewees said that courses were periodically updated to meet changing requirements. Most courses were so modified approximately once a year, once every six months, or at the end of each course. The information indicating that a revision was necessary usually came in the form of feedback from the field, changes in regulations and manuals, suggestions from instructors, and student critiques. The information was fed back to the training managers/instructors through: (1) written reports such as graduate survey questionnaires and field supervisor questionnaires, (2) meetings and conferences, and (3) word of mouth, as might come about through interactions with other instructors and students.

The course review techniques, (as separate from field evaluations) that were mentioned most frequently were student critiques, suggestions from instructors, reviews of field requirements, and supervisor reviews. Opinions of course review techniques, as evoked by a catagorical scale, are presented in Table 15.

BEST COPY AVAILABLE

Table 15

Evaluation of the Course Review Techniques Used
(N = 52)

Classification	% of Respondents
Superior	19
Very good	33
Good	33
Fair	6
Poor	2
No response	7

A very small percentage of the training managers/instructors were dissatisfied with the course review techniques that were used. Only eight per cent rated the course review techniques as "fair" or "poor." The largest proportion of respondents indicated satisfaction ("good," "very good," or "superior") with the techniques.

The principal course review technique problem, as seen by these respondents, was that the reviewers were subjective (as might occur with student critiques and with some instructor suggestions) and in many instances untrained and/or ill prepared to review and evaluate the course. An example of a course review situation being performed by a less than fully prepared training evaluator was provided by an interviewee. In this case, the review was performed by Air Training Command personnel who were not fully aware of specific local problems. The result, according to the interviewee, was a set of suggestions which was unrealistic.

Another course review technique problem mentioned often in the interviews with the training manager/instructors was the problem of lag time. It seems that the time between the reviews and the implementation of the recommendations for revision was too great to be of maximum benefit. Additionally, the training managers/instructors sometimes thought that instructors are overburdened when given the additional responsibility of reviewing course material, new texts, etc., for the purpose of updating a course. These interviewees thought that course reviews should be performed by training evaluation specialists and not instructors.

The training managers/instructors most often indicated that the problem of overtraining can best be identified by student critiques and instructor suggestions. The problem of undertraining, on the other hand, might be identified best by the same two sources in addition to supervisor reports from the field.

BEST COPY AVAILABLE

A critical incident item was used to obtain insights into course review results that were actually used and into the mechanism for such implementation. The interviewees were asked to identify a course review result, the needed revision, who was primarily responsible for bringing about the change, the outcome, and an explanation of how the results were helpful. Almost 40 per cent of the interviewees responded to this item. In every case related, the instructors were responsible for recommending the course revision; in one case a change was introduced to an existing test, in another to course equipment, and in the remainder to course content; in every instance the modification was helpful and beneficial to those concerned. An example of a reported incident follows. The students complained, in the student critiques, that they were having difficulty learning a particular subject matter area. The techniques used in presenting the material to the students were a lecture and slides. The instructors replaced the lecture and slides with a motion picture, in which a step by step procedure was demonstrated. Test results indicated that the problem was solved.

Field Evaluation

Eighty per cent of the respondents claimed that field evaluations were conducted for the purpose of updating the course. For most of the courses involved, the field evaluations were conducted yearly. For some courses, they were conducted semiannually, and for other courses they were conducted quarterly and every other year. The field evaluation techniques consisted of supervisor evaluations of graduates and graduate critiques of the training they received. The field evaluation results were fed back to the training managers/instructors via written reports, meetings, and word of mouth.

As indicated in Table 16, fifty-two per cent of the training manager/instructor sample considered the course review techniques as "superior" or "very good," whereas twenty-seven per cent considered the field evaluations as such. Similarly, 32 per cent considered the field evaluations to be "fair" or "poor," whereas 8 per cent thought the course review techniques were "fair" or "poor" for obtaining information that can be employed for course update.

BEST COPY AVAILABLE

Table 16

Evaluation of the Field Evaluation Techniques Used
(N = 52)

Classification	% of Respondents
Superior	8
Very good	19
Good	17
Fair	17
Poor	15
No response	23

The explanations for the relatively low evaluation were found in the enumeration of the problems with the techniques. The field evaluation problems most frequently mentioned were: subjectivity, careless completion of questionnaires, insufficient samples of field units, and inexperienced evaluators and interviewers conducting field visits. The interviewees suggested that: (1) respondents to the field evaluation questionnaires often comment in areas in which they are unqualified, (2) the field evaluations were not extensive enough to provide valid assessments of graduate performance, (3) field visits should be conducted more frequently and cover all commands, (4) interview teams rather than questionnaires be used because more intensive investigation could be made through this technique, and (5) interviewers that go into the field be trained both in interviewing techniques and in the content area under consideration.

Graduate questionnaires were considered by the training managers/instructors, as for the other groups interviewed, to represent the best technique for determining areas of over and undertraining. Supervisor evaluations were regarded as a good source for identifying areas of undertraining, but not overtraining. Fifty per cent of the sample of training managers/instructors responded to item 28, a critical incident item, which asked the interviewee to relate a specific incident in which a field evaluation result was used by course developers. In all but two of the incidents, so related, an instructor or instructor supervisor implemented the recommended modification. All the incidents, but one, involved modifications to course content. The one incident that did not involve course content was one in which a modification to the existing physical facilities was required. In all cases, the modifications were successful.

BEST COPY AVAILABLE

In one instance the curriculum section was responsible for instituting the change. In this case, a new procedure was being used in the field for loading pallet trains. Since this procedure was not taught in the course, the graduates learned the procedure on the job. As a result of field interviews, the curriculum department learned of the need for updating the course. The request was submitted to the course instructors, who revised the course so that the new procedure was included.

Several of the questions that were included in the interviews with training managers/instructors were also included in the training evaluator's interview. Cross comparisons of the responses of the two groups indicates some opinion differences. For example, a comparison of Tables 6 and 15, which presented the results to the evaluation of the course review techniques that were used in their courses, indicates the responses of the training evaluation specialists to be less favorable to the course review techniques than the responses of the training managers/instructors. About one-third, in each sample, rated the techniques "good." However, while 52 per cent of the training managers/instructors placed the techniques in the "superior" or "very good" categories, only 37 per cent of the training evaluators did so. Similarly, while 8 per cent of the training managers/instructors rated the techniques as "fair" or "poor," 34 per cent of the training evaluators did the same. Since instructors play a more active role in the actual operations of reviewing a course (they were identified as an integral part of a course review technique, namely, instructor suggestions), it is possible that they may be predisposed to view them more favorably. On the other hand this experience with the techniques may place them in an excellent position for providing valid evaluation statements. If merit is assumed for the remarks of the training evaluators, then some modification in course review techniques may be indicated.

A comparison of the assessments of the field evaluation techniques provided by the training manager/instructor and the training evaluation specialties also suggests opinion differences. (Tables 7 and 16). Sixty-two per cent of the training evaluators considered the field evaluation techniques to be "superior" or "very good." However, only 27 per cent of the training managers/instructors evaluated the techniques as "superior" or "very good." None of the training evaluators placed the field evaluation techniques in the fair/poor categorization; 32 per cent of the training managers/instructors considered the field evaluations, as used for updating courses, as "fair" or "poor."

BEST COPY AVAILABLE

The inconsistency across the two different specialties suggests either that each has met with greater success with one technique than with the other or that one group possesses insights (or vested interests) not possessed by the other. If merit is assumed for the remarks of the training managers/instructors, then some modification in certain of the field evaluation techniques may be indicated.

Student Measurement

The various student measurement techniques investigated were: end of block (and unit) written tests, performance tests, criterion checklists, oral tests, instructor ratings, and programmed instruction tests. Passing and failing scores on end of block (and unit) written tests were based on an absolute score. The training managers/instructors thought that the end of block written tests were effective for evaluating students for the same reasons given by the other training specialist groups interviewed. Although the end of block written tests were viewed favorably by many of the interviewees, nevertheless several interviewees (quite correctly) did not think that they are effective for evaluating students who are learning skills (as opposed to learning conceptual material).

Some of the reasons given for the judgment that the written tests were effective were that: (1) the tests include one item each for every enabling and criterion abjective, (2) statistical approaches to test development are used, (3) the tests are periodically reviewed for the purpose of updating, and (4) the tests provide objective evaluations of student achievement. Written tests were said to be practical because: (1) students are familiar with their traditional format, and (2) they isolate the students strong and weak areas.

Policy states that students receive the results of their tests. Students who did not perform satisfactorily on tests either received additional training and/or special assistance and then were retested or repeated the block and then retested. Although students could be dropped from the course (and were in some cases), this was not a popular approach to the problem and was reserved for very special circumstances. Those who were dropped from a course usually had problems that affected their progress in the course (such as physical or emotional problems) but which were not directly related to learning ability.

The training managers/instructors were asked if they were aware of a specific incident in which course developers took corrective action on the basis of student measurement results. Approximately 35 per cent of the interviewees responded. The instructor or course supervisor was mentioned as being responsible for fostering the use of the resulting recommendation. Modifications were made to course content, the equipment, or the tests. In all the incidents related the modifications were found to be beneficial.

BEST COPY AVAILABLE

Student Critique

According to the training managers/instructors, graduates completed cretique forms which were reviewed by the instructors, branch level personnel, department level personnel, and occasionally the instruction and measurement section, curriculum branch, technical school, and center commander.

Sixty-one per cent of the training manager/instructor interviewees related critical incidents in which corrective action was taken on the basis of student suggestions. Instructors, course supervisors, and curriculum section personnel were identified as those who implemented the corrective action. In the incidents related, action was taken in regard to course content, tests, equipment, and course administration. The corrective actions were always found to be reasonable and beneficial.

Evaluation of Instructional Effectiveness

According to the training managers/instructors the criteria that were used to judge instructional effectiveness are the technical competence of students as displayed on their course tests, teaching ability as reviewed by course supervisors, the qualifications of the instructors, adequacy of textual materials, and adequacy and appropriateness of equipment and facilities. The criteria used for judging instructional effectiveness generally did not differ for different courses. Not all the criteria listed above were used in judging instructional effectiveness in every course.

General-Nonspecific Responses

Before the interview was terminated, each interviewee was given the opportunity to state anything he thought would be of some importance to this field survey. The open-ended question, "Is there anything else you would like to tell me?" was asked to obtain this information. The following statements were typical for each of the job specialties sampled and represent current thoughts and attitudes within the various specialties.

BEST COPY AVAILABLE

Instructional System Development Specialists

The ISD method is a superior training system and achieves better training than any other system

ISD is good for trade and skill courses but not necessary for the academic and knowledge courses

There is too much emphasis on cost effectiveness

Increase the number of trained ISD specialists

People at the centers blame the ISD method for the excessive time it takes them to develop a course--the delays are due to their unfamiliarity, not the method

If ISD specialists are not provided with subject matter specialists when developing a course, then provisions should be made for them to take the course

Training Evaluation Specialists

Training evaluation division often is assigned specialists who are not qualified for the position--they often come up through the instructional ranks

Cooperation among the departments is less than good--it is sometimes difficult for course managers to accept the results of training evaluations

Restrict the training evaluation specialists' activities and responsibilities to training evaluation and do not include other aspects of training

Student Measurement Specialists

Explain ISD objectives to course personnel--this could result in time and cost savings in developing courses

The ungraded criterion checklist represents a great improvement--checklists are more appropriate for skill training

Increase standards for graduation from a course--low level graduates can return to the course as instructors and too often they are only mediocre instructors

Preparing courses by the ISD method involves too much work and this has the ultimate effect of reducing instructional effectiveness

BEST COPY AVAILABLE

Training Managers/Instructors

Too much emphasis is placed on student critiques

Criterion checklists are the best thing to come along in testing

ISD is good for low level courses

Standardize on the equipment used in the field so that retraining technicians would not be required so frequently

Field evaluations are poorly conducted--comments on the course content by people who are not recent graduates may not be valid and recent graduates are not sufficiently knowledgable about their jobs to comment on the training they received to prepare them for the job

Importance of ISD Steps

The five ISD steps were evaluated by each group of respondents for the purpose of determining which steps were considered to be most instrumental in achieving an effective instructional program. The rank ordering method was used to obtain this information. In order to avoid positional biases, the five steps were presented in their sequential order (i.e., steps 1, 2, 3, 4, 5) to one portion of the sample and in a randomly based disarray for another portion of the sample (i.e., steps 4, 2, 5, 1, 3).

The instructions to the interviewees were:

> Please assign rank orders, from 1 to 5, in this column so that we can determine which ISD steps you think help most to achieve an effective instructional program. The number one is the best, number 5 represents the worst. If you want to tie two steps, then assign both the average value between the two positions and skip two numbers. For example, if you wanted to rank ISD step 1 and 4 as tied for third position, you would assign the rank of 3.5 to both and give the next ISD step the rank of 5.

Eighty-four per cent of the total sample responded to this item. Mean rankings were obtained for each of the four specialties, as well as for the total sample and are presented as Table 17. Table 17 indicates differences of opinion among the different specialties. However, there was universal agreement that step 5 (conduct and evaluate instruction) was the least important for developing an effective instructional program.

BEST COPY AVAILABLE

Step 1 (analyze system requirements) was seen as the most important step by three of the four specialties. The training evaluators reflected their sphere of interest when they placed step 3 (develop objectives and tests) at the topmost position. Step 1 for the training evaluators was placed in the fourth position. Step 2 (define education and training requirements) maintained a relatively similar position for the four specialties, that is, two groups placed it in second position and two placed it in third position. There were differences of opinion in the ranking of step 3. The training evaluation specialists placed step 3 in the topmost position, the ISD specialists ranked it in the second position, and the student measurement and training manager/instructor personnel placed it fourth. Similarly, step 4 (plan, develop, and validate instruction) was regarded differently among the specialties. The training evaluators ranked it in the second position, to the student measurement and training managers/instructors it was in the third position, and it was ranked fourth by the ISD specialists.

Table 17

Mean Rankings of ISD Steps on Importance in Course Development

ISD (N=33)	TE (N=21)	SM (N=24)	TM/I (N=39)	Total (N-117)
Step 1	Step 3	Step 1	Step 1*	Step 1
Step 3	Step 4	Step 2	Step 2*	Step 2
Step 2	Step 2	Step 4	Step 4	Step 3
Step 4	Step 1	Step 3	Step 3	Step 4
Step 5	Step 5	Step 5	Step 5	Step 5

* tied ranks

When the mean rankings for the total sample are considered, rather than the individual specialties, the order of importance of each step for achieving an effective instructional program was the sequential order of steps 1, 2, 3, 4, 5. Since the ISD method may be considered as a pyramid, with each step resting on the foundation provided by prior steps, this rank ordering may reflect the downstream effects of poor performance of any individual step. From this point of view, poor performance of early steps would impact on more steps than poor performance of later steps.

BEST COPY AVAILABLE

Emphasis on ISD Steps

The ISD steps were also evaluated with reference to the amount of emphasis they received during the development of the course with which each interviewee was most familiar. To this end, the paired comparison method was employed. The interview section titled ISD step comparison, presented as part of Appendix A, shows the method employed for collecting these data. Only 63 per cent of the entire sample provided usable, complete responses. The interviewee had to have knowledge of a course that was developed using the ISD method in order to respond to this item.

The method outlined by Edwards (1957) was used to obtain ISD steps emphasis scale values for each ISD step. The method requires calculation of the frequency with which each step was judged as having been given more emphasis than the other steps. The frequencies were then expressed as proportions (by dividing each frequency by the total number of judgments made). These proportions were corrected to z values (the unit normal deviate scale value which has a true zero point). The result of these calculations is a set of values along a ratio scale with a true zero point. Table 18 presents the results.

Table 18

Scale Values of the Emphasis the ISD Steps Received in Course Development

Step	Scale Value
3	.609
4	.352
5	.165
1	.030
2	.000

Steps 1 and 2 were positioned at the bottom of the emphasis continuum. This finding might be considered to be in line with expectations. In many cases, step 1 (analysis of system requirements) was performed by a team of specialists from headquarters or an existing course was revised in terms of ISD, but the existing system requirements outline was employed. Similarly, step 2 was often in existence before the course was redeveloped. For the remaining three steps, the sequence was steps 3, 4, 5. These findings have implications in terms of development time estimates and the assignment of personnel (i.e., either

BEST COPY AVAILABLE

in terms of ability and experience or man-hours). The finding may also, as for the ranking data discussed above, reflect the feed forward characteristics of the ISD model. We note that the emphasis ordering, at least for the last three steps, parallels the importance ordering discussed above.

In order to gain insight into the internal consistency of the paired comparison data, Kendall's coefficient of consistence (Edwards, 1957) was calculated. Inconsistencies in the paired comparison judgments are indicated by circular triads. A circular triad would exist, for example, if a respondent estimated that step 3 received more emphasis than step 4, and step 4 received more emphasis than step 5, but that step 3 received less emphasis than step 5. The greater the number of circular triads, the more inconsistent was the interviewee in making these estimates.

The coefficient of consistence can vary from 0 to 1.00; 0 represents the maximum number of circular triads, and 1.00 represents the complete absence of circular triads. The coefficients of consistence were determined for each of the respondents. They varied from .2 to 1.00. Of the 87 interviewees who made the paired comparisons of the ISD steps, 22 or 25 per cent had coefficients of consistence less than their judgments. The inconsistencies, as they existed in 25 per cent of the sample, could have been due to an indifference on the part of the interviewee to making careful judgments, to the difficulty in separating two or more very close steps (i.e., close on the continuum of emphasis), or to a personality factor which would predispose one to being inconsistent regardless of the task on hand.

Reliability of Interviewee Responses

Except for the ISD specialists, two pairs of items were included in each interview set in order to check directly on interviewee response consistency. For the ISD specialists, seven pairs of items were so included. Retest measures, as could be obtained in a repeated interview, were not possible. Typically, for each reliability estimation, the two questions were asked at different points in the interview. Although the questions were phrased slightly differently, the expectation was for similarity between the elicited responses. Three different statistical approaches were used to calculate reliability estimates. Where one of the involved variables was continuous in nature and the other variable was reduced to two categories, the biserial correlation coefficient was calculated. This correlation coefficient was obtained for those items where the ISD program was evaluated, once on a qualitative scale ranging from "poor" to "superior" (this scale was collapsed into two categories by grouping the "superior" and "very good" and the "good," "fair," and "poor" responses) and then on a quantitative scale ranging continuously from 0 to 100.

BEST COPY AVAILABLE

A coefficient of contingency was obtained where a correlation index was required between variables, each of which contained more than two classes. This was applicable where the ISD specialists estimated the amount of emphasis devoted to each of the ISD steps in the development of the course and then estimated the relative emphasis given to each step in the paired comparison item, labeled ISD step comparison.

Finally, the percentage of agreement between two items, as in the case where the respondents were asked to suggest ways of improving the use and application of ISD and then, in another item, were asked for the singlemost important improvement. Table 19 presents the items on which reliability estimates were obtained, the type of reliability estimate used, and the obtained reliability estimate.

Table 19

Reliability Estimates of Interviewee Responses

Specialty and Items	Reliability Estimate Used	Obtained Reliability
ISD		
27 and 31	Biserial r	.78
28 and 29	Per cent agreement	64%
4 and Step 1*	Coefficient of contingency	.65
11 and Step 2*	Coefficient of contingency	.61
14 and Step 3*	Coefficient of contingency	.57
17 and Step 4*	Coefficient of contingency	.59
20 and Step 5*	Coefficient of contingency	.52
TE		
18 and 22	Biserial r	.59
19 and 20	Per cent agreement	74%
SM		
38 and 42	Biserial r	.65
39 and 40	Per cent agreement	85%
TM/I		
1 and 44	Biserial r	.55
6 and 42	Per cent agreement	90%

* These data were obtained from the ISD step comparison item.

BEST COPY AVAILABLE

The biserial correlations obtained on the evaluation of the ISD program for achieving its purposes ranged from .55 to .78 (Table 19), the percentage agreement between the items dealing with suggestions for improving the use and application of ISD* ranged from 64 per cent to 90 per cent, and the coefficients of contingency** ranged from .52 to .65.

A rule of thumb has been applied to the minimal adequate level of reliability in the interview situation (Krech & Crutchfield, 1948, pp. 260-262). The minimal adequate level, by this rule of thumb, for group measurement is a correlation coefficient .50. Since none of the correlation estimates obtained in this survey were below .50, and since the percentage agreement data represent correlation coefficients which are considerably in excess of this value, it appears that the interviewees were adequately consistent in their attitudes and opinions.

ISD Evaluation as a Function of Location and Experience

The assessment of the ISD program was examined from two additional, but diverse, points of view. An analysis of variance was performed to determine if the quantitative (on the scale of 0 to 100) evaluation of ISD varied as a function of: (1) Air Force base, and (2) length of time in the job assignment. The inclusion of the second variable was based on the conjecture that respondents who had been on their present job longest would be more resistent to change and, therefore, would give lower evaluations of the ISD system.

Four ATC technical schools were used in this analysis: Lackland, Sheppard, Chanute, and Lowry. Keesler Technical Training Center was omitted because of an insufficient sample size. Two levels of experience on present assignment were included. Level I consisted of respondents with less than two years in their current assignment; Level II consisted of incumbents with more than five years in the current assignment. A sample of five respondents was randomly selected from each of the four bases, in each experience level. The Sheppard AFB Level II cell had only four ISD evaluations. The fifth evaluation was approximated by taking the arithmetic mean of the four evaluations. Thus, every cell in the analysis of variance had five entries. Table 20 presents a summary of the results of the variance analysis.

---

* The method employed for determining the percentage of agreement involved inspecting the responses to the item in which an unlimited number of suggestions could be made for improving ISD for a match with the item in which the one most important improvement for ISD was identified. A match was considered to exist if the most important improvement suggested also appeared among the general improvements suggested.

** In calculating the coefficients of contingency, the number of pluses that each respondent assigned to the ISD step, as summed in the vertical column of the ISD step comparison matrix, represented one variable (the greater the number of pluses the more emphasis the step received). The scale value (0 to 100) represented the other variable.

BEST COPY AVAILABLE

Table 20

Summary of Analysis of Variance for Location and Experience

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F
Location	269.49	3	89.83	.20
Experience	773.16	1	77 .16	1.71
Location x Experience	836.99	3	279.00	.62
Error	14553.75	32	451.68	
Total	16333.39	39		

The variance analysis indicated that no statistically significant difference existed between experience levels or among locations. The location by experience interaction was, similarly, not statistically signficant. Hence, these data provide no support for a contention suggesting a different evaluation of the ISD system as a function of ATC bases or time on current job assignment.

BEST COPY AVAILABLE

CHAPTER IV

DISCUSSION

The prior chapters presented the results from the field interviews conducted with the ISD, student evaluation, training manager/instructor, and course evaluation specialist at five ATC centers. The present chapter attempts to integrate some of the more salient findings into a logical structure and to present some of the implications of the findings.

Evaluation of ISD

Each of the four groups interviewed seemed to think that the ISD procedure had a different value in terms of achieving its general purpose of graduating students who meet job performance requirements. For comparative purposes, the mean magnitude estimate, for each group, of the extent to which the ISD procedure achieves its goal is presented in Figure 5.

Regardless of the individual mean numerical value assigned by the various groups, each group can be said to have regarded the ISD procedures favorably. This conclusion is also supported by the responses of the various groups when they were asked to make a categorical judgment of the extent to which the ISD methods achieve their purpose.

From the point of view of the general and overall effectiveness of the ISD procedures, it seems that they have proven their value, at least for the various groups involved in their implementation. Each of the various groups has had exposure to different aspects of the ISD procedures and, accordingly, can be anticipated to have a somewhat different reaction.

BEST COPY AVAILABLE

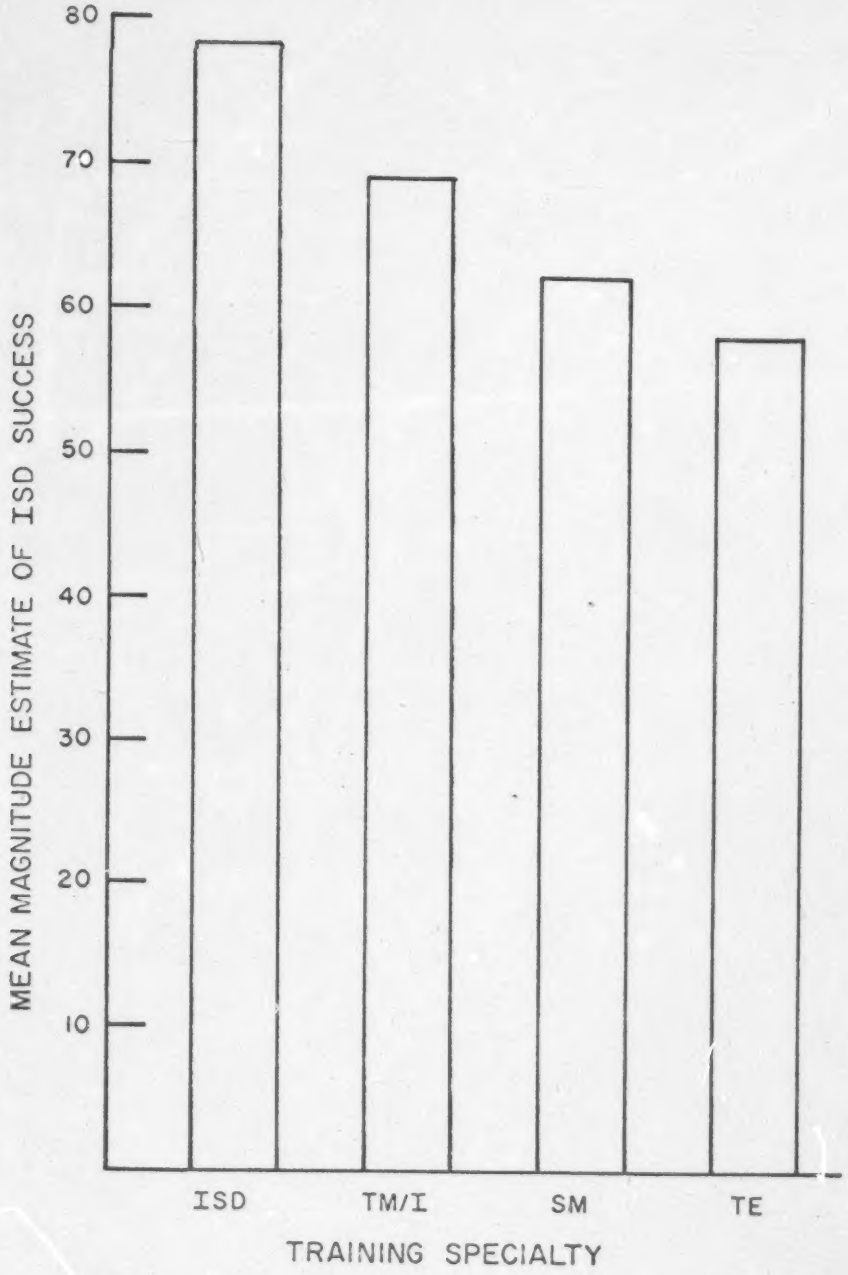


Figure 5. Mean magnitude estimates by training specialty (0-100) of ISD for achieving its goals.

BEST COPY AVAILABLE

Problems Within the ISD Methods

Regardless of the general and overall positive reaction to the ISD methods manifested by the groups involved, a number of problems within the use and implementation of ISD became evident.

First, there was a prevailing overcurrent of misconception about ISD and the need for an information dissemination program to clarify these misconceptions. Some of the information gaps were based on faulty information, e.g., "ISD will remove the instructor from the system and all training will be programmed." Others were less obvious and were related to a general lack of insight on the part of some supervisory personnel regarding the advantages of and the methods involved in the ISD procedures. While the ISD procedures have been presented in several formal documents, these documents have evidently not achieved an indoctrinational purpose at the supervisory level. Accordingly, there is a need for a more general and extensive method for orienting supervisory personnel in the attributes and goals of the ISD system.

Effective ISD implementation is also dependent on an appropriately trained corps of personnel who are thoroughly prepared in the performance of their individual assignments and who possess adequate insight into the specific aspects of other ISD requirements. It is believed that, in some instances, ISD aspects may have been resisted because the implications of the work for subsequent ISD steps were not fully perceived. Accordingly, some priority seems indicated for attendance at the ISD course, especially for those currently concerned in a direct manner with ISD application.

It is also possible that with increased training (and experience), the time required to apply the ISD procedures will decrease. Certainly, the interviewees bemoaned what they considered to be the excessive time involved in applying the various ISD procedures. This problem is interactive with the frequently suggested need for more personnel. It is also quite possible that the perceived need for an augmentation in the number of personnel assigned to performance of the ISD steps will diminish with increased proficiency in the techniques involved.

However, we note that there is no readily accessible body of data regarding the amount of time required to complete the various ISD steps for various types of courses. Such data will be required before definitive statements can be made regarding the time required to complete the ISD process for any course.

BEST COPY AVAILABLE

The need for "streamlined" ISD procedures was also mentioned. It is possible that some of the time involved performing the various ISD steps could be shortened through the provision of a variety of job aids, work methods reorganization, organizational modification, and an increase in communications efficiency. Which one or ones, if any, of these would assist most in reducing the time consumed in applying the various ISD steps cannot be known as a result of the present study. It is most probable that the various ISD steps would benefit differentially from the various modifications. Certainly, investigation of the feasibility of such solutions should take place before personnel allowances are augmented.

On the other hand, if the increased training and other suggestions, presented above, for reducing the time required to apply the ISD methods to a course are not tenable, then the interviewee suggested increase in the manpower allocation seems reasonable if high priority is to be placed on application of the ISD method to all courses. Such universal application seems supportable in view of the reported degree of success for the ISD procedures wherever they have been applied.

We note that the ISD model is often rejected by instructors who are teaching a course developed by some other model and are satisfied with the course and the end products it produces. It seems that some value may exist in a concept which advocates introducing ISD to the training centers first for those courses that have an immediate need for revision.

There is also a need for the evaluation of how well each of the ISD steps is performed. There are several available analytic techniques that could be used for the various steps. Evaluation methods that might be considered for evaluating the various steps of the ISD model are: comprehensiveness and relevance checks, statistical methods, interviews and judgments, rating scales, experimental methods, checklists, written tests, performance tests, and task analytic matrix methods.* It is likely that suggestions for modifications to the application and utilization of ISD would result from evaluations. The introduction of the modifications to the model would serve to make the ISD procedures more useful and acceptable to application personnel.

Finally, we note that some users have elected to implement ISD through performance of a selected subset of the ISD system steps. Since each ISD step is dovetailed with a prior step, such abbreviation does not seem tenable. It seems that the user should apply the total system. Short cutting certain required steps may serve to undermine the total system.

---

* These techniques are discussed fully in Siegel, A.I., Bergman, B.A., Federman, P., & Sellman, W.S. Handbook for the Evaluation of Technical Training Courses and Students (AFHRL-TR-72-15).

BEST COPY AVAILABLE

Student Measurement Techniques

The interviewees seemed quite satisfied, on the whole, with the student measurement techniques employed. The criterion referenced checklist was especially believed to be valuable.

However, there was little evidence that the reliability of the widely employed end of block written tests is determined for each test employed. Some of these tests seem quite short and, accordingly, could approach or be lower than the lowest acceptable reliability limit.

Another end of block written test consideration is the frequency with which such tests are updated. It seems that updates as frequent as the end of each course may be too frequent, with the exception of the most obvious sorts of modification. On the other hand, updating a test about once a year may be too infrequent. Guidelines are required which will allow satisfaction of the purposes of updating a test without so long a delay as to render the test outmoded or so short a delay as to require update on the basis of an insufficient data substrate.

In some of the interviews, the item analytic data were reviewed. It was noted that although the required item analyses were conducted, the tests contained items which, at least by the usual psychometric standards, are not acceptable. Accordingly, it seems that item acceptability psychometric guidelines are also necessary. Such guidelines are also contained in the above referenced Handbook (AFHRL-TR-72-15).

We are aware that the time allowed for student measurement within a course is limited. This limitation reemphasizes the need for powerful test devices, where they are used. On the other hand, there seems to be little gained by constricting the testing time to the point that the testing is not effective for achieving its purpose. Such a procedure represents a waste of time. Poor testing may, in fact, be worse than no testing at all.

There also seems to be a need for alternate or parallel forms of end of block written tests for use in situations in which a student who fails originally is recycled and retested.

Additionally, there was little evidence regarding the predictive validity of these tests. Nor was evidence suggested regarding their correlation with various selection and classification tests. The development of such data seems mandatory. Ongoing research into new and improved testing methods also seems indicated.

BEST COPY AVAILABLE

Finally, there seemed to be little reliance on the confidence testing and the hierarchical testing approaches which have come under some degree of acceptance in other contexts. Application of these concepts, at least on a trial basis, seems to be a tenable recommendation.

Course Evaluation Techniques

The principal course evaluation technique employed was "critiques." These took the form of student critiques and critiques by graduates and their supervisors. These latter critiques are performed after the graduate has been on his technical job for a period of time. Sometimes the critique data are collected through questionnaire methods; less frequently, interviews are conducted. The questionnaire methods are economical and seem to produce valuable insights into areas for needed course revision. They neither seem to restrict unduly the range or freedom of response of the respondent nor do they demand an unusual or excessive amount of time for completion.

There are some comment, however, that the methods employed for these critiques lack objectivity. Additionally, the interpretation of the data yielded by these methods was reported to be often performed by persons who are not completely familiar with the total subject matter content or with the context in which their recommendations must be implemented.

Since the course evaluation constitutes the quality control aspect of the ISD system, a need for considerable research seems to exist in this sphere. Although there is a considerable body of techniques available and although there is little, if any, doubt that some objectivity could be built into the already existing critiques by ad hoc methods, such fixes seem, at best, temporary. Accordingly, we suggest that a carefully designed research program be instituted in the course techniques arena. Such a program would build on the already extant state of the knowledge in the course evaluation sphere and develop these, for Air Force application, so that the course review techniques can yield longitudinal, objective, quantitative, continuous course review information. Such techniques, in addition to being fully synergistic with the ISD system, must take advantage of the most advanced knowledge in posttraining performance measurement and meet acceptable criterion requirements, e.g., reliability, objectivity, scorability, analyzability, practicality, predictability, etc. Such a program would also possess considerable implications for selection and classification methodology and validation.

BEST COPY AVAILABLE

At any rate, the use of untrained course evaluation personnel and of untrained interviewers (as reported) for performing course evaluation does not seem to constitute an acceptable practice. Interviewer training is relatively easy to accomplish and at least reliability, if not validity, can be built into the interview by such training. It also seems that such interviews should be at least semistandardized and fully criterion referenced. (We note that criterion referencing of the various critique forms may also be indicated.)

Finally, there was comment that the time lag between course reviews and the implementation of the recommendations they elicit is often lengthy. This suggests a lack of direction or standardization regarding course update. It may also suggest a lack of coordination between course review personnel and those persons who are responsible for course update. Since course changes are implemented mostly by instructor personnel, close coordination between these two groups seems manadatory if course modifications are to be implemented effectively.

Postscript

Any new program, such as the ISD program, must undergo monitoring and a series of refinements and calibrations before it achieves its maximum potential. The results of the present study suggest that the ISD program has achieved considerable success, and support its continued development. The emphasis of the current set of interviews was on the student measurement and the course review techniques employed within the ISD program. The suggestions regarding modification and/or refinement of certain aspects of the methods currently employed are aimed at improving the power of these methods. The performance evaluation technology is highly dynamic, and it is highly likely that this aspect of the ISD system will require ongoing modification in accordance with ongoing developments in the field.

BEST COPY AVAILABLE

REFERENCES

Bergman, A.A., & Siegel, A.I. Training evaluation and student measurement: A review of the literature. Denver: Lowry Air Force Base, AFHRL-TR-72-3, AD 747040.

Edwards, A.I. Techniques of attitude scale construction. New York: Appleton-Century-Crofts, 1957.

Krech, D., & Crutchfield, R.S. Theory and problems of social psychology. New York: McGraw-Hill, 1948.

Siegel, A.I., Bergman, B.A., & Federman, P. Some techniques for the evaluation of technical training courses and students. Denver: Lowry Air Force Base, AFHRL-TR-72-15, AD 753094.

Torgerson, W.S. Theory and methods of scaling. New York: Wiley, 1958.

BEST COPY AVAILABLE

APPENDIX A

BEST COPY AVAILABLE

INQUIRY INTO INSTRUCTIONAL SYSTEM DEVELOPMENT
TECHNIQUES EMPLOYMENT

Instructional System Development

BEST COPY AVAILABLE

My name is _____, I represent Applied Psychological Services. We are under contract with the Air Force Human Resources Research Laboratory (AFSC) to develop a comprehensive handbook of training evaluation and student measurement methods. The interviews we are conducting represent a part of the study involved in the development of this handbook.

The purpose of the interview is to learn about the training evaluation and student measurement techniques being used at the various Air Training Command schools. In addition, we are interested in learning about the Instructional System Development (ISD) program and its effectiveness in the design of Air Force technical courses. The data and insights gained from the interviews will provide important information for the subsequent development of the handbook for training evaluation and student measurement.

Your cooperation in the interview is greatly appreciated and will represent a major contribution to the study. Of course, everything you say will be used for the purposes of the study only and will remain completely anonymous. None of the information we collect will be attributed to any specific person.

BEST COPY AVAILABLE

INTRODUCTION

(HAND INTERVIEWEE SET OF DEFINITIONS AND ISD FLOW CHART AND ALLOW A FEW MINUTES FOR HIM TO READ THEM.)

1. (HAND INTERVIEWEE CARD ISD-1) Would you say that you are "familiar," "moderately familiar," or "unfamiliar" with the Instructional System Development program?

(IF INTERVIEWEE IS "FAMILIAR" OR "MODERATELY FAMILIAR," ASK QUESTIONS 2 AND 3. IF INTERVIEWEE IS "UNFAMILIAR" WITH ISD PROGRAM DISCONTINUE INTERVIEW.)

2. What was the one source from which you gained most of your familiarity with the ISD program?

3. (HAND INTERVIEWEE CARD ISD-3) Would you describe your depth of experience with ISD as "extensive," "moderate," or "slight"?

(IF INTERVIEWEE FITS ONE OF THE X-ED CELLS BELOW, CONTINUE WITH QUESTION 4. IF INTERVIEWEE HAS ONLY SLIGHT EXPERIENCE WITH ISD DISCONTINUE INTERVIEW.)

QUESTION 1

	Familiar	Moderately Familiar	Unfamiliar
Extensive	X	X	
Moderate	X	X	
Slight			

SPECIFIC APPLICATION

I. ANALYZE SYSTEM REQUIREMENTS

(SAY TO INTERVIEWEE: THE NEXT SET OF QUESTIONS WILL DEAL SPECIFICALLY WITH EACH OF THE FIVE ISD STEPS. IN ANSWERING THESE QUESTIONS THINK OF THE SAME COURSE YOU JUST IDENTIFIED AS BEING MOST FAMILIAR WITH.)

4. (HAND INTERVIEWEE CARD ISD-4) In terms of a scale which ranges from 0 to 100, how much emphasis would you say was placed on the ISD step, Analyze System Requirements, in planning the training program?

5. Why so?

BEST COPY AVAILABLE

(IF ISD WAS EMPHASIZED WITH A VALUE > 20 ASK QUESTIONS 6, 7, AND 8; OTHERWISE GO TO QUESTIONS 9 AND 10.)

6. Using the same scale of values, 0 to 100, how successful was the analysis of system requirements step in helping to plan the course content?

7. Why so?

8. How was the ISD step, Analyze System Requirements, helpful in planning the training program?

9. (IF ISD WAS NOT EMPHASIZED OR EMPHASIZED VERY LITTLE, ASK) Why wasn't the ISD step, Analyze System Requirements, emphasized (or emphasized more)?

10. Can you recall a specific situation when the ISD principle, Analyze System Requirements, was used and it did not really help in planning the training program? Tell me the incident, the need, who fostered the use, the outcome, and why it did not help so much.

II. DEFINE EDUCATION OR TRAINING REQUIREMENTS

11. (HAND INTERVIEWEE CARD ISD-4) Using the scale values again, how much emphasis was placed on the ISD step, Define Education or Training Requirements, in planning the training program?

12. Why so?

13. (IF ISD WAS NOT EMPHASIZED OR EMPHASIZED VERY LITTLE, ASK) Why wasn't the ISD step, Define Education or Training Requirements, emphasized (or emphasized more)?

III. DEVELOP OBJECTIVES AND TEST

14. (HAND INTERVIEWEE CARD ISD-4) How much emphasis was placed on the ISD step, Develop Objectives and Tests, in planning the training program?

15. Why so?

16. (IF ISD WAS NOT EMPHASIZED OR EMPHASIZED VERY LITTLE, ASK) Why wasn't the ISD step, Develop Objectives and Tests, emphasized (or emphasized more)?

BEST COPY AVAILABLE

IV. PLAN, DEVELOP, AND VALIDATE INSTRUCTION

17. (HAND INTERVIEWEE CARD ISD-4) How much emphasis was placed on the ISD principal, Plan, Develop, and Validate Instruction, in developing and implementing the course?

18. Why so?

19. (IF ISD WAS NOT EMPHASIZED OR EMPHASIZED VERY LITTLE, ASK) Why wasn't the ISD step, Plan, Develop, and Validate Instruction, emphasized (or emphasized more)?

V. CONDUCT AND EVALUATE INSTRUCTION

20. (HAND INTERVIEWEE CARD ISD-4) Using the scale values, 0 to 100, what was the extent to which the step, Conduct and Evaluate Instruction, was emphasized in implementing and controlling the effectiveness of the course?

21. Why so?

(IF ISD WAS EMPHASIZED WITH A VALUE > 20, ASK QUESTIONS 22, 23, and 24; OTHERWISE GO TO QUESTIONS 25 AND 26.)

22. Using the scale again, how successful is the ISD step, Conduct and Evaluate Instruction, for implementing and controlling the effectiveness of the course?

23. Why so?

24. How was this ISD step of conducting and evaluating instruction helpful in implementing and controlling the effectiveness of the course?

25. (IF ISD WAS NOT EMPHASIZED OR EMPHASIZED VERY LITTLE, ASK) Why wasn't this ISD step emphasized (or emphasized more)?

26. Can you think of a specific situation when the ISD step, Conduct and Evaluate Instruction, was applied and it did not prove to be beneficial in implementing and controlling the course? Tell me the incident, the need, who fostered the use, the outcome, and why it did not help so much.

BEST COPY AVAILABLE

GENERAL OPINION

27. In general, would you rate or describe the ISD program as "superior," "very good," "good," "fair," or "poor" for achieving the stated objectives of turning out graduates who meet job performance requirements efficiently and effectively?

28. What specific or general suggestions do you have for improving the use and application of ISD?

29. What one improvement would help ISD utilization most?

30. In your opinion, why do some people avoid ISD utilization?

31. (HAND INTERVIEWEE CARD ISD-31) How would you evaluate the ISD program, on the basis of values 0 to 100, for achieving its purposes?

32. (HAND INTERVIEWEE ANSWER SHEET) Please assign rank orders, from 1 to 5, in this column (POINT) so that we can determine which ISD steps you think help most to achieve an effective instructional program. The number one is the best, number 5 represents the worst. If you want to tie two steps, then assign both the average value between those two positions and skip two numbers. For example, if you wanted to rank ISD steps 1 and 4 as tied for third position, you would assign the rank of 3.5 to both and give the next ISD step the rank of 5.

33. What do you think are the three greatest problems in using the ISD model?

34. Is there anything else you would like to tell me?

BEST COPY AVAILABLE

Familiar: Know and understand all of the ISD steps

Moderately familiar: Know and understand most, but not all, of the ISD steps

Unfamiliar: Do not know any of the ISD steps

Handout Card for Item 1

Extensive: Used and applied all the ISD steps

Moderate: Used and applied most, but not all, of the ISD steps

Slight: Never used or applied any of the ISD steps

Handout Card for Item 3

Scale of Emphasis

0	10	20	30	40	50	60	70	80	90	100
Very Little				Moderate				Extensive		

Handout Card for Item 4

Scale of Achievement

0	10	20	30	40	50	60	70	80	90	100
Poor			Fair	Good			Very Good		Superior	

Handout Card for Item 31

BEST COPY AVAILABLE

INQUIRY INTO INSTRUCTIONAL SYSTEM DEVELOPMENT
TECHNIQUES EMPLOYMENT

Training Evaluation

My name is _____, I represent Applied Psychological Services. We are under contract with the Air Force Human Resources Research Laboratory (AFSC) to develop a comprehensive handbook of training evaluation and student measurement methods. The interviews we are conducting represent a part of the study involved in the development of this handbook.

The purpose of the interview is to learn about the training evaluation and student measurement techniques being used at the various Air Training Command schools. In addition, we are interested in learning about the Instructional System Development (ISD) program and its effectiveness in the design of Air Force technical courses. The data and insights gained from the interviews will provide important information for the subsequent development of the handbook for training evaluation and student measurement.

Your cooperation in the interview is greatly appreciated and will represent a major contribution to the study. Of course, everything you say will be used for the purposes of the study only and will remain completely anonymous. None of the information we collect will be attributed to any specific person.

BEST COPY AVAILABLE

(SAY TO INTERVIEWEE: WE DEFINE TRAINING EVALUATION AS THE DETERMINATION OF HOW WELL LEARNING OBJECTIVES WERE IDENTIFIED, THE EFFICIENCY OF THE ISD PROCESS, AND THE PROFICIENCY OF GRADUATES IN PERFORMING THE JOBS FOR WHICH THEY WERE TRAINED. SUCH EVALUATION OFTEN LEADS TO COURSE REVISION. ANSWER THESE QUESTIONS IN TERMS OF THE COURSE WITH WHICH YOU SAID YOU ARE MOST FAMILIAR.)

1. Was the course periodically updated to meet changing requirements?

2. (IF YES) How often?

3. (IF YES) Where did the information indicating a needed course revision come from?

4. (IF YES) How did the information get fed back to you?

5. What sorts of course review techniques were used, as separate from field evaluation?

6. Would you rate the course review techniques (as separate from field evaluation) used for obtaining information that can be used for modifying the course as "superior," "very good," "good," "fair," or "poor?"

7. What was the principal problem with the course review techniques employed?

8. Which course review techniques best identified areas of over and under training?

9. Tell me about a specific incident in which course review results were actually used? Tell me the incident, the need, who fostered the use, the outcome, and why it was helpful.

10. Were field evaluations conducted for the purpose of updating the course contents?

11. (IF YES) How often?

12. (IF YES) What types of field evaluation techniques are (were) used?

13. (IF YES) How did this need get fed back to you?

14. Would you rate the field evaluation techniques used for obtaining information that can be used for modifying the course as "superior," "very good," "good," "fair," or "poor?"

BEST COPY AVAILABLE

15. What was the principal problem with the field evaluation techniques employed?

16. Which field evaluation technique best identified areas of over and under training?

17. Tell me about a specific incident in which the field evaluation results were actually used by the course developers. Tell me the incident, the need, who fostered the use, the outcome, and why it was helpful.

18. In general, would you rate or describe the ISD program as "superior," "very good," "good," "fair," or "poor" for achieving the stated objectives of turning out graduates who meet job performance requirements efficiently and effectively?

19. What one improvement would help ISD utilization most?

20. What specific or general suggestions do you have for improving the use and application of ISD?

21. In your opinion, why do some people avoid ISD utilization?

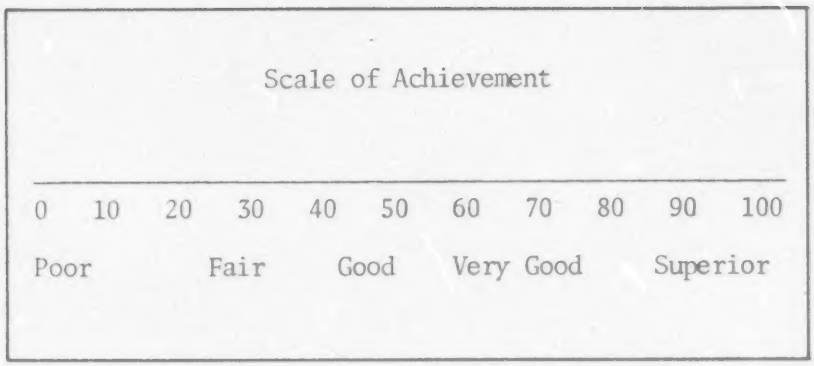
22. (HAND INTERVIEWEE CARD TE-22) How would you evaluate the ISD program, on the basis of the values 0 to 100, for achieving its purposes?

23. (HAND INTERVIEWEE ANSWER SHEET) Please assign rank orders, from 1 to 5, in this column (POINT) so that we can determine which ISD steps you think help most to achieve an effective instructional program. The number one is the best, number 5 represents the worst. If you want to tie two steps, then assign both the average value between those two positions and skip two numbers. For example, if you wanted to rank ISD steps 1 and 4 as tied for third position, you would assign the rank of 3.5 to both and give the next ISD step the rank of 5.

24. What do you think are the three greatest problems in using the ISD model?

25. Is there anything else you would like to tell me?

BEST COPY AVAILABLE



Handout Card for Item 22

BEST COPY AVAILABLE

INQUIRY INTO INSTRUCTIONAL SYSTEM DEVELOPMENT
TECHNIQUES EMPLOYMENT

Student Measurement

BEST COPY AVAILABLE

My name is _____, I represent Applied Psychological Services. We are under contract with the Air Force Human Resources Research Laboratory (AFSC) to develop a comprehensive handbook of training evaluation and student measurement methods. The interviews we are conducting represent a part of the study involved in the development of this handbook.

The purpose of the interview is to learn about the training evaluation and student measurement techniques being used at the various Air Training Command schools. In addition, we are interested in learning about the Instructional System Development (ISD) program and its effectiveness in the design of Air Force technical courses. The data and insights gained from the interviews will provide important information for the subsequent development of the handbook for training evaluation and student measurement.

Your cooperation in the interview is greatly appreciated and will represent a major contribution to the study. Of course, everything you say will be used for the purposes of the study only and will remain completely anonymous. None of the information we collect will be attributed to any specific person.

BEST COPY AVAILABLE

(SAY TO INTERVIEWEE: WE HAVE SEVERAL QUESTIONS ON STUDENT MEASUREMENT TECHNIQUES. WE DEFINE STUDENT MEASUREMENT AS MEASURES FOR THE PURPOSE OF EVALUATING INDIVIDUAL STUDENT ACHIEVEMENT AND PROGRESS. ANSWER THESE QUESTIONS IN TERMS OF THE COURSE WITH WHICH YOU SAID YOU ARE MOST FAMILIAR.)

1. (SHOW CARD SM-1) What are (were) the different types of student measurement (student achievement) techniques employed?

2. On the basis of end of block written test results, how did you know if a student passed or failed or how well he did?

(SAY TO INTERVIEWEE: THINK ABOUT THE METHODS EMPLOYED FOR CONSTRUCTING END OF BLOCK WRITTEN TESTS).

3. Was a formal item analysis conducted?

4. (IF YES) How so?

5. Was any attempt made to correlate the end of block test scores with other measures of performance?

6. (IF YES) How so?

7. Why do you think that these end of block written tests were effective for evaluating students?

8. How so?

9. What was the principal problem, in your opinion, in using these written tests?

10. How so?

11. Were the students informed of their end of block written test results?

12. (IF YES) What recommendations were made to those students who do not achieve satisfactory results?

13. Were work sample tests employed?

(IF YES, ASK QUESTIONS 14, 15, 16, 17, 18, 19, and 20; OTHERWISE GO TO QUESTION 21)

14. How did you know, on the basis of the work sample test scores, whether or not a student passed or failed or how well he did?

BEST COPY AVAILABLE

BEST COPY AVAILABLE

15. Why do you think work sample tests were effective for measuring student achievement?

16. How so?

17. What was the principal problem, in your opinion, with the work sample tests?

18. How so?

19. Were students informed of their work sample test results?

20. (IF YES) What recommendations were made to those students who did not achieve satisfactory results?

21. Were criterion checklists employed?

(IF YES, ASK QUESTIONS 22, 23, 24, 25, 26, 27, and 28; OTHERWISE GO TO QUESTION 29)

22. On the basis of the criterion checklist, how did you know if a student passed or failed or how well he did?

23. Why do you think criterion checklists were effective for measuring student achievment?

24. How so?

25. What was the principal problem, in your opinion, with the criterion checklists?

26. How so?

27. Were students informed of their criterion checklist results?

28 (IF YES) What recommendations were made to those students who did not achieve satisfactory results?

29. Were there any student measurement techniques not used in the course that you would have like to see used? If so, what are they?

30. How were the results of student evaluations fed back to the course developers?

31. How were the student evaluations used by the course developers?

BEST COPY AVAILABLE

32. Can you tell me of a specific incident where the course developers took some correction action on the basis of student measurement results? Tell me the incident, the need, who fostered the use, the outcome, and why it was helpful.

33. How frequently were student measurement techniques reviewed?

34. How frequently were student measurement techniques modified as a result of these reviews?

35. Did the students fill out critique forms on the course?

36. (IF YES) Who reviewed these critiques?

37. Can you tell me a specific incident where the course developers took some corrective action on the basis of the critique forms filled out by the students? Tell me the incident, need, who fostered the use, the outcome, and why it was helpful.

38. In general, would you rate or describe the ISD program as "superior," "very good," "good," "fair," or "poor" for achieving the stated objectives of turning out graduates who meet job performance requirements efficiently and effectively?

39. What specific or general suggestions do you have for improving the use and application of ISD?

40. What one improvement would help ISD utilization most?

41. In your opinion, why do some people avoid ISD utilization?

42. (HAND INTERVIEWEE CARD SM-42) How would you evaluate the ISD program, on the basis of the values 0 to 100, for achieving its purposes?

43. (HAND INTERVIEWEE ANSWER SHEET) Please assign rank orders, from 1 to 5, in this column (POINT) so that we can determine which ISD steps you think help most to achieve an effective instructional program. The number one is the best, number 5 represents the worst. If you want to tie two steps, then assign both the average value between those two positions and skip two numbers. For example, if you wanted to rank ISD steps 1 and 4 as tied for third position, you would assign the rank of 3.5 to both and give the next ISD step the rank of 5.

44. What do you think are the three greatest problems in using the ISD model?

45. Is there anything else you would like to tell me?

BEST COPY AVAILABLE

BEST COPY AVAILABLE

End of block written test
Criterion Checklist
Programmed instruction test
Instructor's rating
Oral test
Work sample performance test

Handout Card for Item 1

Scale of Achievement

0	10	20	30	40	50	60	70	80	90	100
Poor			Fair		Good		Very Good		Superior	

Handout Card for Item 42

BEST COPY AVAILABLE

INQUIRY INTO INSTRUCTIONAL SYSTEM DEVELOPMENT
TECHNIQUES EMPLOYMENT

Training Managers/Instructors

BEST COPY AVAILABLE

My name is _____, I represent Applied Psychological Services. We are under contract with the Air Force Human Resources Research Laboratory (AFSC) to develop a comprehensive handbook of training evaluation and student measurement methods. The interviews we are conducting represent a part of the study involved in the development of this handbook.

The purpose of the interview is to learn about the training evaluation and student measurement techniques being used at the various Air Training Command schools. In addition, we are interested in learning about the Instructional System Development (ISD) program and its effectiveness in the design of Air Force technical courses. The data and insights gained from the interviews will provide important information for the subsequent development of the handbook for training evaluation and student measurement.

Your cooperation in the interview is greatly appreciated and will represent a major contribution to the study. Of course, everything you say will be used for the purposes of the study only and will remain completely anonymous. None of the information we collect will be attributed to any specific person.

BEST COPY AVAILABLE

1. In general, would you rate or describe the ISD program as "superior," "very good," "good," "fair," or "poor" for achieving the stated objectives of turning out graduates who meet job performance requirements efficiently and effectively?

2. (HAND INTERVIEWEE CARD TM/I-2) I would like you to select from the adjectives that appear on this card the three that you think best describe the ISD model.

3. (HAND INTERVIEWEE CARD TM/I-3) Now I would like you to do the same with this list of adjectives.

4. What would you say were the three best aspects of the ISD model? Rank order these, e.g., 1 = best, 3 = least best.

5. What would you say were the three worst aspects of the ISD model? Rank order these, e.g., 1 = worst, 3 = least worst.

6. What changes would you like to see introduced that would improve the ISD program?

7. Why do you support the employment of ISD?

8. What kinds of problems do course developers typically find when developing new ISD courses?

9. What sorts of criteria do you have for judging instructional effectiveness?

10. Do these criteria differ for different courses?

11. (IF YES) How so?

(SAY TO INTERVIEWEE: NOW WE HAVE SOME QUESTIONS ON TRAINING EVALUATION. WE DEFINE TRAINING EVALUATION AS THE DETERMINATION OF HOW WELL LEARNING OBJECTIVES WERE IDENTIFIED, THE EFFICIENCY OF THE ISD PROCESS, AND THE PROFICIENCY OF GRADUATES IN PERFORMING THE JOBS FOR WHICH THEY WERE TRAINED. SUCH EVALUATION OFTEN LEADS TO COURSE REVISION. ANSWER THESE QUESTIONS IN TERMS OF THE COURSE WITH WHICH YOU SAID YOU ARE MOST FAMILIAR).

12. Was the course periodically updated to meet changing requirements?

13. (IF YES) How often?

14. (IF YES) Where did the information indicating a needed course revision come from?

15. (IF YES) How did the information get fed back to you?

BEST COPY AVAILABLE

BEST COPY AVAILABLE

16. What sorts of course review techniques were used, as separate from field evaluation?

17. Would you rate the course review techniques (as separate from field evaluation) used for obtaining information that can be used for modifying the course as "superior," "very good," "good," "fair," or "poor?"

18. What was the principal problem with the course review techniques employed?

19. Which course review technique best identified areas of over and under training?

20. Tell me about a specific incident in which the course review results were actually used. Tell me the incident, the need, who fostered the use, the outcome, and why it helped so much.

21. Were field evaluations conducted for the purpose of updating the course contents?

22. (IF YES) How often?

23. (IF YES) What types of field evaluation techniques are (were) used?

24. (IF YES) How did this need get fed back to you?

25. (IF YES) Would you rate the field evaluation techniques used for obtaining information that can be used for modifying the course as "superior," "very good," "good," "fair," or "poor?"

26. (IF YES) What was the principal problem with the field evaluation techniques employed?

27. (IF YES) Which field evaluation technique best identified areas of over and under training?

28. (IF YES) Tell me about a specific incident in which the field evaluation results were actually used by the course developers. Tell me the incident, the need, who fostered the use, the outcome, and why it helped so much.

(SAY TO INTERVIEWEE: WE HAVE SEVERAL QUESTIONS ON STUDENT MEASUREMENT TECHNIQUES. WE DEFINE STUDENT MEASUREMENT AS MEASURES FOR THE PURPOSE OF EVALUATING INDIVIDUAL STUDENT ACHIEVEMENT AND PROGRESS. ANSWER THESE QUESTIONS IN TERMS OF THE SAME COURSE WE WERE DISCUSSING.)

BEST COPY AVAILABLE

29. What are (were) the different types of student measurement (student achievement) techniques employed?

30. On the basis of end of block written test results, how did you know if a student passed or failed or how well he did?

31. Why do you think that these end of block written tests were effective for evaluating students?

32. How so?

33. Are the students informed of their end of block written test results?

34. (IF YES) What recommendations are made to those students who do not get satisfactory evaluations?

35. How were the results of student measurements fed back to the course developers?

36. How were the student measurements used by the course developers?

37. Can you tell me of a specific incident where the course developers took some corrective action on the basis of student measurement results? Tell me the incident, the need, who fostered the use, the outcome, and why it was so helpful.

38. Did the students fill out critique forms on the course?

39. (IF YES) Who reviewed these critiques?

40. Can you tell me a specific incident where the course developers took some corrective action on the basis of the critique forms filled out by the students? Tell me the incident, the need, who fostered the use, the outcome, and why it was so helpful.

41. What specific or general suggestions do you have for improving the use and application of ISD?

42. What one improvement would help ISD utilization most?

43. In your opinion, why do some people avoid ISD utilization?

44. (HAND INTERVIEWEE CARD TM/I-44) How would you evaluate the ISD program on the basis of the values of 0 to 100 for achieving its purposes?

BEST COPY AVAILABLE

45. (HAND INTERVIEWEE ANSWER SHEET) Please assign rank orders, from 1 to 5 in this column (POINT) so that we can determine which ISD steps help most to achieve an effective instructional program. The number one is the best, number 5 represents the worst. If you want to tie two steps, then assign both the average value between those two positions and skip two numbers. For example, if you wanted to rank ISD steps 1 and 4 as tied for third position, you would assign the rank of 3.5 to both and give the next ISD step the rank of 5.

46. What procedures have been instituted to evaluate the worth of ISD?

47. Which aspects of ISD do you like best? Why?

48. Which aspects of ISD do you like least? Why?

49. Is there anything else you would like to tell me?

BEST COPY AVAILABLE

helpful
appropriate
complete
good
profitable
organized
efficient
excellent
effective
dynamic

Handout Card for Item 2

mundane
unworkable
poor
static
unsuccessful
weak
limited
unsuitable
tedious
costly

Handout Card for Item 3

Scale of Achievement

0	10	20	30	40	50	60	70	80	90	100
Poor			Fair		Good			Very Good		Superior

Handout Card for Item 44

BEST COPY AVAILABLE

BEST COPY AVAILABLE

INQUIRY INTO INSTRUCTIONAL SYSTEM DEVELOPMENT
TECHNIQUES EMPLOYMENT

Headquarters

BEST COPY AVAILABLE

1. In general, would you rate or describe the ISD program as "superior," "very good," "good," "fair," or "poor" for achieving the stated objectives of turning out graduates who meet job performance requirements efficiently and effectively?

2. What specific suggestions do you have for improving the use and application of ISD?

3. What one improvement would help ISD utilization most?

4. In your opinion, why do some people avoid ISD utilization?

5. (HAND INTERVIEWEE CARD HQ-5) I would like you to select from the adjectives that appear on this card the three that you think best describes the ISD model?

6. (HAND INTERVIEWEE CARD HQ-6) Now I would like you to do the same with this list of adjectives.

7. What would you say were the three best aspects of the ISD model? Rank order these, i.e., 1 = best, 3 = least best.

8. What would you say were the three worst aspects of the ISD model? Rank order these, i.e., 1 = worst, 3 = least worst.

9. What changes would you like to see introduced that would improve the ISD program?

10. Why do you support the employment of ISD?

11. What kinds of problems do course developers typically find when developing new ISD courses?

12. (HAND INTERVIEWEE CARD HQ-12) How would you evaluate the ISD program on the basis of the values of 0 to 100, for achieiving its purposes?

13. What procedures have been instituted to evaluate the worth of ISD?

14. What sorts of criteria does headquarters have for judging instructional effectiveness?

15. Do these criteria differ for different courses?

16. (IF YES) How so?

17. Which aspects of ISD do you like best? Why?

18. Which aspects of ISD do you like least? Why?

19. Do you have any plans or desires to terminate the usage of ISD?

20. (IF YES) How so?

21. What are the three most significant problems headquarters staff personnel have in using ISD?

22. Is there anything you would like to tell me?

helpful
appropriate
complete
good
profitable
organized
efficient
excellent
effective
dynamic

Handout Card for Item 5

mundane
unworkable
poor
static
unsuccessful
weak
limited
unsuitable
tedious
costly

Handout Card for Item 6

Scale of Achievement

0	10	20	30	40	50	60	70	80	90	100
Poor			Fair			Good		Very Good		Superior

Handout Card for Item 12

BEST COPY AVAILABLE

ISD STEP COMPARISONS

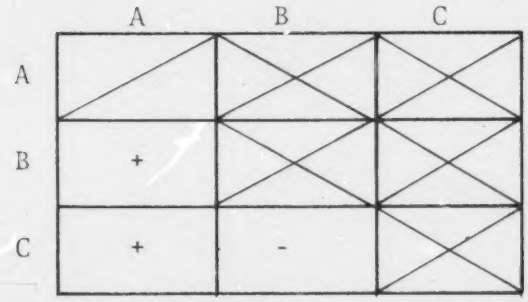
The purpose of this rating form is to determine how much each ISD step was emphasized (relative to every other step) in achieving the ultimate purpose of graduating students who meet job performance requirements. The five ISD steps to be rated for emphasis in achieving the purposes of the course are arranged around the chart below. Thinking only about the same course that was identified in the interview as the one with which you were most familiar, start by comparing ISD step 1 (across) with ISD step 2 (down). If you think that step 1 (across) was emphasized more than step 2 (down), place a plus sign (+) in step 1's column (alongside step 2). If you think that step 1 (across) was emphasized less than step 2 (down), place a minus sign (-) in step 1's column (alongside step 2).

In the example below, three different instructional methods were rated for meeting the objectives of the course. The objectives were designated by the letters A, B, and C. The judge, starting by comparing A with B, decided that A was more effective than B. Thus, he placed a plus sign in A's column (alongside B). In the next judgment, he compared A with C and decided that A was more effective than C. A plus sign was thus placed in A's column (alongside C). In comparing B with C, the judge decided that B was less effective than C and therefore a minus sign in B's column (alongside C).

BEST COPY AVAILABLE

Example

Course Objectives



In completing the chart below, work down each column by comparing step 1 with steps 2, then step 3, through step 5. Next, work down column 2 by comparing step 2 with step 3 through 5. Continue the process of comparing pairs of ISD steps until the chart is completed (a total of 10 comparisons).

ISD STEP COMPARISONS

	STEP 1 ANALYZE SYSTEM REQUIREMENTS	STEP 2 DEFINE EDUCATION OR TRAINING REQUIREMENTS	STEP 3 DEVELOP OBJECTIVES AND TESTS	STEP 4 PLAN, DEVELOP, AND VALIDATE INSTRUCTION	STEP 5 CONDUCT AND EVALUATE INSTRUCTION
ANALYZE SYSTEM REQUIREMENTS	X	X	X	X	X
DEFINE EDUCATION OR TRAINING REQUIREMENTS		X	X	X	X
DEVELOP OBJECTIVES AND TESTS			X	X	X
PLAN, DEVELOP, AND VALIDATE INSTRUCTION				X	X
CONDUCT AND EVALUATE INSTRUCTION					X

Name _____

BEST COPY AVAILABLE

BEST COPY AVAILABLE

RESPONSE SHEETS

Training Evaluation

Interviewee's Name _____ Location _____

Current Assignment _____ How Long _____ Date _____

ISD Course and Number Most Familiar With _____
(CROSS OUT "ISD" IF COURSE DID NOT HAVE ISD MODEL APPLIED TO IT)

1. ____yes ____no

2. ____at the end of each course ____quarterly ____semi-annually

____yearly ____ every other year ____ other (specify)_____

3. _____

4. ____written report____meeting____word of mouth____other (specify)____

_____

5. ____student interviews____instructor interviews ____instructor suggestions

____review of field requirements ____0ther (specify)_____

6. ____superior ____very good ____good____fair ____poor

7. _____
_____
_____
_____
_____
_____

8. _____
_____
_____
_____
_____

BEST COPY AVAILABLE

9. Incident _____

Need _____

Press _____

Outcome _____

Why helpful _____

10. ____ yes _____ no

11. _____at the end of each course _____quarterly _____semi-annually ____yearly

_____ once every other year _____other (specify)_____

12. _____graduate evaluations _____supervisor's evaluations of graduates

_____graduate's critique of training received _____other (specify)_____

_____

13. _____written report _____meeting _____word of mouth _____Other (specify)__

_____

14. _____superior _____very good _____good _____fair _____poor

15. _____

_____

_____

_____

_____

16. _____

_____

_____

_____

17. Incident_____

Need_____

Press_____

Outcome_____

Why helpful_____

BEST COPY AVAILABLE

18. ____superior ____very good ____good ____fair ____poor

19. _____

_____

_____

_____

20. (R-19)_____

_____

_____

21. _____

_____

_____

_____

22. (R-18) _____

23. (R) ISD Steps — Rank Order

Plan, Develop, and Validate Instruction	_____
Define Education or Training Requirements	_____
Conduct and Evaluate Instruction	_____
Analyze System Requirements	_____
Develop Objectives and Tests	_____

24. (1) _____

(2) _____

(3) _____

25. _____

_____

_____

_____

_____